

48/550

FIGURE 48

MGALIISGSSAGPVTQASLPPWGLSHGRCGFLLYMENTLCSHRTQSFSELSQSLMRPGFLQM
PYISCAKLSKIWFPAASKPCLLAFLEVELLSRSLFSKMICFLFLSFLFPPHIYTHAS

Important features of the protein:

Signal peptide:

amino acids 1-41

Transmembrane domain:

amino acids 88-107

Casein kinase II phosphorylation site.

amino acids 47-50

N-myristoylation site.

amino acids 24-29

49/550

FIGURE 49

GGCTTCTACAGTCCACAACACCCACCAGCCCCAGGCCAGCAGAATGAGCCAGTGAGTGGCGGGGCTCCAGTT
TGGCTGTTGCTATGACAACGTGGCCACTGCAGCCGGTCTCTTGGGGAAGGCTGTGTGGGCCAGCCAGCCATGC
CTACCCCGTGCGGTGCCTGCTGCCAGTGCCCATGGCTCTTGTGCAGACTGGGCTGCCGCTGGTACTTCGTTGC
CTCTGTGGGCCAATGTAACCGCTTCTGGTATGGCGGCTGCCATGGCAATGCCAATAACTTTGCCTCGGAGCAAGA
GTGCATGAGCAGCTGCCAGGGATCTCTCCATGGGCCCGTCTGCCAGCCTGGGGCTTCTGGAAGGAGCACCCA
CACGGATGGTGGCGGCAGCAGTCTTGCAGGCGAGCAGGAACCCAGCCAGCACAGGACAGGGGCCGCGGTGCAGAG
AAAGCCCTGGCCCTTCTGGTGGTCTCTGGCGGCAAGACCAACAGCCTGGGCCAGGGGAGGCCCCCACACCCAGGC
CTTTGGAGAATGGCCATGGGGGCGAGGAGCTTGGGTCCAGGGCCCCCTGGACTGGGTGGAGATGCCGGATCACCAGC
GCCACCCCTTCCACAGCTCCTCTACAGATCTCAGTTCCACCTCTCCAGGATTAGCTTGGCAGGTGTGGAGCCCT
CGTTGGTGCAGGACGCCCTGGGGCAGTTGGTGGCGGCTCTCCTGCTCAGACGACACTGCCCGGAATCCCAGGCTG
CCTGGCAGAAAGATGGCCAGCCCATCTCCTCTGACAGGCACAGGCTGCAGTTCGACGGATCCCTGATCATCCACC
CCCTGCAGGCAGAGGACGCGGGCACCTACAGCTGTGGCAGCACCCGGCCAGGCCGCGACTCCCAGAAGATCCAAC
TCCGCATTATAGGGGGTGACATGGCCGTGCTGTCTGAGGCTGAGCTGAGCCGCTTCCCTCAGCCCAGGGACCCAG
CTCAGGACTTTGGCCAAGCGGGGGCTGCTGGGCCCTGGGGGCCATCCCCTCTTACACCCACAGCCTGCAACA
GGCTGCGTTTGGACCAGAACAGCCCCGGGTGGTGGATGCCAGTCCAGGCCAGCGGATCCGGATGACCTGCCGPG
CCGAAGGCTTCCCGCCCCCAGCCATCGAGTGGCAGAGAGATGGGCAGCCTGTCTCTTCTCCAGACACCAGCTGC
AGCCTGATGGCTCCCTGGTCATTAGCCGAGTGGCTGTAGAAGATGGCGGCTTCTACACCTGTGTGCTTCAATG
GGCAGGACCGAGACCAGCGATGGGTCCAGCTCAGAGTTCTGGGGGAGCTGACAACTCAGGACTGCCCTCAGTGT
TGACAGTCCAGAGGGGTGATACCGCCAGGCTATTGTGTGTGGTAGCAGGAGAAAGTGTGAACATCAGGTGGTCCA
GGAACGGGCTACCTGTGCAGGCTGATGGCCACCGTGTCCACCAGTCCCAGATGGCAGCTGCTCATTACAACT
TGCGGGCCAGGGATGAGGGCTCCTACATGTGCAGTGCCTACCAGGGGAGCCAGGCAGTCAGCCGACGACCCGAGG
TGAAGGTGGTCTCACCAGCACCCACCGCCAGCCAGGGACCCTGGCAGGGAAGTGGCTCGACCAGCCAGAGCTGG
CCAAGTGTGATTTGATCTTGCAGGCCAGCTTTGTGGCAATGAGTATTACTCCAGCTTCTGCTGTGCCAGCTGTT
CACGTTTCCAGCCTCACGCTCAGCCCATCTGGCAGTAGGGATGAAGGCTAGTTCCAGCCCCAGTCCAAAATAGTT
CATAGGGCTAGGGAGAAAGAGATGGACTCTTGGCTTCTCTCTCTGGCTGGCAAAGGGAGTTATCTTCTGGAA
TACATTAGCTCTTTCAAAAACCCACCCAGTGTTTAGCCTCAACGGCAGCCAGTTACCAGCTTCTCTCTGTAGCCT
TCAGCAGTGTGTCATCTCTGACATAACCACAGGCTGCTGTTTTCAAGAAGAGCAATCTGTTTGGATAAGAAAAA
CCTTTACTTTACAGCTTCCCTTTATAATTTGTTACACAGGAATAGTTAATGCATTTGTTTGTGTTTTTGTGAG
ACGGAGTTTCACTCTTGTGTCGCCAGGCTGGAGGGCAATGGCGCGATCTCAGCTCACTGCAACCTCCGTCTCCTGG
GTTCTTGATTCTCCTGTGTACGCTTCTGAGTAGCTGGGATTACAGATGCCTATCACCATGCCTGGGTAAATTTTT
GTATTTTTTAGTTGAGATGGGGTTTCGCCATGTTGGCCAGGCTGGTCTCGAAGTCTTGACCTCAGATGATCTGCC
GCCTCAGCCTCCCAAAGTGTCTGGGATFACAGGCATGAGCCACCAGCCAGCCATCAATGCATTTTTTTTATTTT
TTTTTTGAGACAGAGTTTCGCACTTCTTGGCCAGGCTGGAGTACAATGGTGGGATCTTGGCTCACTGCAACCTCC
ACCTCCTGGGTTCAAGCGCTTCTCCAGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGTATGTGCCACCATGCCT
GGCTAATTTTTGTATTTTTTGGTGGAGACGGGGTTTCTCCATGTTGGTCAGACTGGTCTTGAAGTCCCAGCCTCAGG
TAATCCGCCCCGCTCCGCTCCCAAATGCTGGGATTAGAGGTGTGAGCCACTGTGCCAGCCCATCAATGTGTT
TTAAAGCTAGCTGTGAGGGTTCCACTTAATTTAAAGCTGGGCAGGGAGATGTGTAATGATTTCAAAGTTAACACC
TGTTTGTGTTTCAAAGGGCATGCCAAGTCTGCTGTATCAGGGAAGTATCTGTGCTAAATCAGCGATGGTTCA
TTGCTCTAGTCTCTCTCACCCTTCTAGGCAGTGCATCAGTCAGCTCTAAATCTGGTGCAGAGGGTTAACAGCATA
ACCTTGTTGGCAAAATGGAATAGATGTTAAGACCTCAATAGGGATTTGGGATGAAACAGCTGCAGTTAGCACT
GTTATCTGAGCATGAAAGAACTGGAAACGCTCCTTACGTCGAGATGTTGGACCTTGAAGCCCTCCTGAGGCCAAC
ATGCAAACTCTGGCTGTGACGGTTCTCTGACACCTGTGTAAAGCTGACCAGCCTGCTCTGTACAGTGACAATGAG
GAGCCCCCTCTCTTCTTAAGTAGGAATCTGTGAAGCAAAATGTTTGTGCAAGACAAATCAGACTGTCACTCA
TTAAAAACAGCATTAGCAGGATGAGGATAGCAATGGGAAGGGTTGTGGGCAATGCAGTAACAGGGAAATGGCTT
CAGAAATGGTTTGTGAGTTGGAAGACAACTTCTCATCTCTCAGGACTTCTAATTCCTTGATGCTAAAAGAAGAGG
CATGGATTCTATGAGCTTCCAAGTCCCTTTCCACTTTAACTTCTACAAATCTTTCAGAGGACTGCCTAGTAGCA
AAGGTTATTCTGGACACAGGAAAGACGGGCATFACAGGGACCAAGCTCTGAAAGGTGACTTTTATTACCAACA
CACTGGCTGGAAAAGGGACAAACCACATCACGGGTGAGTGATACTTCTCAGTCTTCTCTACTCATTCAACAAAGG
AAATGTGGGCTGGGGCAGAGCTCTTTTTTCATTTAATACGGAAAAATATTGAAGAGCATCCATGTTCACTTAAG
GCTGGTTTTGCTATAGAAATTGGAATAAAGGCCACTTTTTTG

50/550

FIGURE 50

MGPVVPSLGLLEGAPTRMVAAAVLQASRNPASTGQGPRCRES PGLLVVSGGKTNSLGQGRPPT
PRPLENGHGGRSLGPGPLDWVEMPDHQRHPSTAPPTDLTSHLSRISLAGVEPSLVQAALGQLV
RLSCSDDTAPESQAAWQKDGQPISSDRHRLQFDGSLIIHPLQAEDAGTYSCGSTRPGRDSQKI
QLRIIGDMAVLSEAELSRFPQPRDPAQDFGQAGAAGPLGAIPSSHPOQANRLRLDQNPQPRVV
DASPGQRIRMTTCRAEGFPPPAIEWQRDGGQFVSSPRHQLQPDGSLVISRVAVEDGGFYTCVAFN
GQDRDQRWVQLRVLGELTISGLPPTVTVPPEGDTARLLCVVAGESVNIRWSRNGLPVQADGHRV
HQSPDGTLLIYNLRARDEGSYMCSAYQGSQAVSRSTEVKVVSPAPTAQPRDEGRDCVDQPELA
NCDLILQAQLCGNEYYSFCCASC SRFQPHAQPIWQ

Important features of the protein:**Signal peptide:**

amino acids 1-16

Tyrosine kinase phosphorylation site.

amino acids 392-400

N-myristoylation sites.amino acids 9-15, 50-56, 112-118, 146-152, 173-179, 195-201,
220-226, 229-235, 280-286, 306-312, 336-342, 397-403**Myelin P0 protein.**

amino acids 153-182

51/550

FIGURE 51

CAGGCAGAAGCGAACAAAGACCCAGCAAGAGAAGGCAGAGGCTAAGACCCATCCCGTATCTGC
TCTCCTGAAATAATTCTGGAGTCA**ATG**CCTGAAATGCCAGAGGACATGGAGCAGGAGGAAGTTA
ACATCCCTAATAGGAGGGTTCTGGTTACTGGTGCCACTGGGCTTCTTGGCAGAGCTGTACACA
AAGAATTTTCAGCAGAATAATTGGCATGCAGTTGGCTGTGGTTTCAGAAGAGCAAGACCAAAT
TTGAACAGGTTAATCTGTTGGATTCTAATGCAGTTCATCACATCATTTCATGATTTTCAGCCCC
ATGTTATAGTACATTGTGCAGCAGAGAGAAGACCAGATGTTGTAGAAAATCAGCCAGATGCTG
CCTCTCAACTTAATGTGGATGCTTCTGGGAATTTAGCAAAGGAAGCAGCTGCTGTTGGAGCAT
TTCTCATCTACATTAGCTCAGATTATGTATTTGATGGAACAAATCCACCTTACAGAGAGGAAG
ACATACCAGCTCCCCTAAATTTGTATGGCAAAACAAAATTAGATGGAGAAAAGGCTGTCCTGG
AGAACAATCTAGGAGCTGCTGTTTTGAGGATTCCTATTCTGTATGGGGAAGTTGAAAAGCTCG
AAGAAAGTGCTGTGACTGTTATGTTTGATAAAGTGCAGTTCAGCAACAAGTCAGCAAACATGG
ATCACTGGCAGCAGAGGTTCCCCACACATGTCAAAGATGTGGCCACTGTGTGCCGGCAGCTAG
CAGAGAAGAGAATGCTGGATCCATCAATTAAGGGAACCTTTCCTGCTGCTGGCAATGAACAGA
TGACTAAGTATGAAATGGCATGTGCAATTGCAGATGCCTTCAACCTCCCCAGCAGTCACTTAA
GACCTATTACTGACAGCCCTGTCCTAGGAGCACACGTCGGAGAAATGCTCAGCTTGACTGCT
CCAAATTGGAGACCTTGGGCATTGGCCAACGAACACCATTTCGAATTGGAATCAAAGAATCAC
TTTGGCCTTTCCCTCATTGACAAGAGATGGAGACAAACGGTCTTTTCAT**TAG**TTTATTTGTGTTG
GGTTCTTTTTTTTTTTTAAATGAAAAGTATAGTATGTGGCACTTTTTTAAAGAACAAAGGAAATA
GTTTTGTATGAGTACTTTAATTGTGACTCTTAGGATCTTTCAGGTAAATGATGCTCTTGCACT
AGTGAAATTGTCTAAAGAACTAAAGGGCAGTCATGCCCTGTTTGAGTAATTTTTCTTTTTTA
TCATTTTGTGTCTGGCTAAACTTGGAGTTTGAGTATAGTAAATTATGATCCTTAAATATT
TGAGAGTCAGGATGAAGCAGATCTGCTGTAGACTTTTCAGATGAAATTGTTTCATTCTCGTAAC
CTCCATATTTTCAGGATTTTTGAAGCTGTTGACCTTTTCATGTTGATTATTTTAAATTGTGTG
AAATAGTATAAAAATCATTGGTGTTTCATTATTTGCTTTGCCTGAGCTCAGATCAAATGTTTG
AAGAAAGGAACCTTTATTTTTGCAAGTTACGTACAGTTTTTATGCTTGAGATATTTCAACATGT
TATGTATATTGGAACCTCTACAGCTTGATGCCTCCTGCTTTTATAGCAGTTTATGGGGAGCAC
TTGAAAGAGCGTGTGTACATGTATTTTTTTTCTAGGCAAACATTGAATGCAAACGTGTATTTT
TTTAATATAAATATATAACTGTCCTTTTCATCCCATGTTGCCGCTAAGTGATATTTTCATATGT
GTGGTTATACTCATAATAATGGGCCTTGTAAGTCTTTTCACCATTTCATGAATAATAATAATA
TGTAAGTGTGGCATGTAATGCTTAGTTTTCTGTATTTACTTCTTTTTTTTAAATGTAAGGACC
AACTTCTAAACTAATTGTTCTTTTTGTTGCTTTAATTTTTTAAAATTACATTCTTCTGATGTA
ACATGTGATACATACAAAAGAATATAGTTTAATATGTATTGAAATAAAACACAATAAAATT

52/550

FIGURE 52

MPEMPEDMEQEEVNIPNRRVLVTGATGLLGRAVHKEFQQNNWHAVGCGFRRARPKFEQVNLLD
SNAVHHIIHDFQPHVIVHCAAERRPDVVENQPDAAASQLNVDASGNLAKEAAVGAFLIYISSD
YVFDGTNPPYREEDIAPPLNLYGKTKLDGEKAVLENNLGA AVLRIPILYGEVEKLEESAVTVM
FDKVQFSNKSANMDHWQQREFPTHVKDVATVCRQLAEKRMLDPSIKGTFHWSGNEQMTKYEMAC
AIADAFNLPSSHLRPITDSPVLGAQRPRNAQLDCSKLETGLGIGQRTPFIRIGIKESLWPFLLDK
RWRQTVFH

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 105-127

N-glycosylation site.

amino acids 197-201

N-myristoylation site.

amino acids 303-309

Short-chain dehydrogenases/reductases family proteins.

amino acids 18-30

53/550

FIGURE 53

TGGGCTCCCTCCAGCACTGCTGTTGCGCTGCTGCGCTAAGATGGGTGACACTTGGGCCCAGCTTCCCTGGCCTGGGC
CACCCACCCAGCAATGCTGCTGATCTCCCTCCCTCTTGGGAGCCGGGTTGATGCACTCGGATGCCGGCACCAGCT
GCCCCGTCTTTTGCACATGCCGTAACCAGGTGGTGGATTGTAGCAGCCAGCGGCTATTCTCCGTGCCCCAGACC
TGCCAATGGACACCCGAAACCTCAGCCTGGCCCCACAACCGCATCACAGCAGTGGCGCCTGGCTACCTCACATGCT
ACATGGAGCTCCAGGTGCTGGATTGTCACAACAACCTCCTTAATGGAGCTGCCCCGGGGCCTCTTCTCCATGCCA
AGCGCTTGGCACACTTGGACCTGAGCTACAACAATTTTCAAGCATGTGCCAGCCGACATGTTCCAGGAGGCCCATG
GGCTAGTCCACATCGACCTGAGCCACAACCCCTGGCTGGGGAGGGTGCATCCCCAGGCCTTTTCAAGGCCTCATGC
AGCTCCGAGACCTGGACCTCAGTTATGGGGGCTGGCCTTCCCTCAGCCTGGAGGCTCTTGAGGGCCTACCGGGC
TGGTGACCCCTGCAGATCGGTGGCAATCCCTGGGTGTGTGGCTGCACCATGGAACCCCTGCTGAAGTGGCTGCGAA
ACCGGATCCAGCGCTGTACAGCAGATTCTCAGCTGGCTGAGTGCCTGGGGGCCCTCCTGAAGTGGAGGGCGCCCCG
TCTTCTCACTCACTGAGGAGAGCTTCAAGGCCCTGCCACCTGACCCTGACCCTGGATGATTACCTATTTCATTGCGT
TCGTGGGCTTCGTGGTCTCCATTGCTTCTGTGGCCACCAACTTCTCCTGGGCATCACTGCCAACTGCTGCCACC
GCTGGAGCAAGGCCAGTGAAGAGGAAGAGATCTGACATGCCTGCCTCTCATCCCTCCATGCTGCTGACCGCCACA
GCTGCTGGCCACCAGACGCCCTCCCTGATTGCTCACTCTGGTTCCATGGTGACCTGGCTGCCTCAGTCATGGTTC
AAGCAAGGTGGGACACTCATTTTGTATGAGCATCTGCTTTGGGGCCAGGCGGCACGCTAGGAATTGGGAACATCA
GATGAAGTCACTCAGTCCCTGCCCTCAAGGCACCTTCCCTCTGGTCAAGGAGAGAGATCCAAAACCTATTCCCTTT
AAGACTATATGTCAAGACTCTGAGCACGTCATTATGGAGGCCAGAGGAGGAGCCATCATCTGTATCTAGCAATG
TCCATGAGAATTATAAGATTAGAGTGATTGTGAAGTGGGTGATCAGGAAATATCTACTTTGTCAAGTAGGCAAA
GAAGGTGTCTGCACATGGCAGAGGCCAGAATATGCATAGTGTGCTGTGTTGAGAAGAGTGAACAGTTTCTGTCT
ACTTACTTGTATAGAGGGGTGTGGCACAGAACTCAAAACCTACCCCTCACCTCCTGACACCAAACTGTCAAGCTC
TCAGCAATGCCAGCCACTGCCTACAGGGAGTAAGAACACCTCTATGACAGCCCTGGCCTCCTTCCACCAGCAGC
TACCAGGTGAGACCACCTCCAGTGACTGCCCCCATATGACCAATGTCAACAGTTGGTGAGGTCCCAGGCAGCA
GGCTGAGGATGGACACTTTCAATGCCCTTGCTCCTGCCTCTCACTCAAGTTTTGCTTCAGAAGAGAGAGGCAGGA
GGCCAGCAACTGGGGCAGCAAGAGTCTGGCACCTTGGGATCCTAATCATGTGACTGTTCTTGCCACAGTGCTC
ATGCCACAGGGTCTCACCAGGAAAGTGCCTGTGGGCCACAGACCCACAGCCTGGCAGCACCCAGAGCTAAAAGG
GGACAAAGGCAGCACAGTTATGACCATATGAGGCTTTGCATTTTCTTCTAAGCAACTTACCCACGTTAAGCATGA
GGGTGAGAGAGCTATTAAATACTAAGCCCTGGCAGTGTCAAGTACTTTGAAAAGCTCTTGCACAAACCATTC
CTTTGACACACACACACACAAATCTTTGAGGTGAACGCTGTTGTTGTTCCCATTTTACGGATGAGCAACTAAGGCT
CAGAGAGGTTAAAGTCACATGCCACTATGAGCAAGATAAAGTCTGTGCTCTTCTACTGCCCATCCAAGTTGGG
GAACATCACCATTCCCTCTAGAGTTATATAAATTCAAATTCAACTAGAGCTGACAAAGTTCCCTCATAAGGTCCAG
GCACTCCTCTGGGCACTTTTATATCTATTGACTCACTTCTTTCAATTCTCACAGCAACACTGCCTGGTGCTTTT
ATTATCCCCATTTGACAGATGAATTAATCCTAGAGAGTGGAGTGACTTAACCAAGGTTGTCTGGATAAGCCCTAG
AAGGAAGGCGGTAGGCAGCTCCATTCAGGGAACTGCATCTAATCAGTCAGTCAAAATCAAGTAACCTTACGAG
CAAAGCACAAATTATCATCATCGTGGTCTTCTTCATCAGTTTCGTCAAGCAGCATCATTATCTTCCCTCTATTTGT
CAGCACCGGATAGTTTATGAGTATTTTGCATCATTCTCCTTGACTTTTACATCCCTGTGCAGGAGGTAAGTCA
AACATCAGTAATCTGTTTACAGATGGGGAAGGCTCTCAAGTTGGATATGACTTGTCTTGGCAAGGTTG
GGGCTCAACCCCTAACACAGTTCCTTTCCAGTGCTTTCTCAAGTGCTTGGGGAAGAGAATGCCTCAGAAGGCTGG
GTAGTGGGGCCTGGAATTCAAGCATCCATGAATGTGCTAGTGGATAAGCTAAATAGAAGGCAGCCAAACCCATCT
GCTGTACAGATTGAACATATGCTCACGGTAGGGCAAAATGCAGGCTCTGAACAGAGACTACACAGGTAACACCTG
AATAGGAGACTCCTGCTTTACAATGTGTAGATAAAACATCAGCAATGGTGGCCATGGTGGCAGTCATGTGAAAAG
TAAGATCTTTGGGAATCAAGAAAGGAAGCTGTGTTAACCACCTCCTGCTCAAGCCCTGCTGCGTGTGTTGCAAGAG
ATACTAAGAGAGCAAGAAAGCTATAGGTGAGAACTCTGCAGTTTAGGAGAAGAACATCAAGGCACAGTCCAACA
TGCTGATAAGTCTGGCCAGGAGGAGAATTAACACAGGGGCTTTCCACACCTCCCTTGCCCCAAGCTCCAGCGGTA
TTCTATCAGCCCATCCTCCTGGAAAGCCTGAAAGGAATGAAGGAGGCTAATAAGTCATCTTCCAGGAAGGCATCC
CTCACTCGTGCTTCCCTGAGCTAGTCAACCAAAAGAGTCTTCAGAAACCTTGTAGACCTGAAGTACTTGAACCT
GTGTCCCTGAATCTTTCTTACAACATCTGGGACAAATCCCTGGTCTGTGACATCCGAAGCAGAAGTGTGCCCT
GCTCTCTCCTTCTGTGATGACCAAGGATGGTGAACCTCAAGTTGTTCTCTACAAGCCAGGCCAGCAACCTAAATAC
TTGGAGAGGAACCTTTAGAACTATAATCCTGACAAATAGAAAAGTTTCCCATAGGGGCATACCATAATACTAT
AATAACCTCCAGGAACCTATTGTTTGGCAAAATCTAGTTAATATATTTTAAGATATATGCTTTTGTGATAGGAC
TAGAACCAGAAAAGACACCAATGCCCCCTTGACATCAATGTCTTTCTAGTGGGACAATTTGGTCTCCATTAAT
GCCAAACCTTTCTGAACAGGATACATGGCTTTTAAAGGACAGATGTTTCTCCTGCTGCTAGAAGTTCCCTCAGTTT
ACTAGAGCACAAATGAGGAAAGTATTCACCTCCCTACTGCCAAGGAATCCCTGCTTCTCCCCACCGCCATCAT
CTTGTCCAAGCTATCAGAAGCAACCTTCTAGAGATAATCTAACAATCCTGATTAGAATTGCTCCCATATCCCTGG
TGACCACAGGCTTCACTCAAAATGTGCCAACTGGTTAACATGTATGTGATGGGGTATCTCTGCATCTGTATGTCT
GTCTGCGAGGTTCTTGTATATTTGGCTGTCCGCTGACTTGGGACAGATCTCTCTAGAAGTTGGGTTCAAGTTCTCT
GACATAGTCCACTCAGCCATAGGCTGAGTGGCTAAATATGCATAAATAAGCATGCCTAATAGGCATATATAGGT
TGGTGCAAAAGTAATTGCGGTTTTTGGCATTAAAAATGATGGCAAAATCCCAATTACTTTTGGCTCAATCTAAT
ATTACATTGCTTGATAGATTAGATGGAATCCACCAGGTTTAGGGTAGGACTGGATGCTCAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

54/550

FIGURE 54

MLLISLLLAAGLMHSDAGTSCPVLCTCRNQVVDCCSSQRLFSVPPDLPMDTRNLSLAHNRRITAV
PPGYLTCYMELQVLDLHNNSLMELPRGLFLHAKRLAHLDSLNNFHVFPADMFQEAHGLVHID
LSHNPWLRRVHPQAFQGLMQLRDLDSLGGGLAFLSLEALEGLPGLVTLQIGGNPWVCGCTMEF
LLKWLNRNIQRCTADSQLAECRGPPEVEGAPLFSLTEESFKACHLTTLDDYLFIAFVGVVVS
IASVATNELLGITANCCHRWSKASEEEEI

Important features of the protein:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 241-260

N-glycosylation sites.

amino acids 52-55, 81-84, 107-110

Tyrosine kinase phosphorylation site.

amino acids 148-154

N-myristoylation sites.

amino acids 11-15, 263-268

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 175-185

Leucine zipper pattern.

amino acids 77-98

55/550

FIGURE 55

GGCTGCGCCCAGGCCGGCGGGCCCCAGCAGCTGCGAACC GCCGGCGCACCCACCTGTTTCCGCGC
CCGGGGACTTCCCCGGCGGGGCTCAGAAGTGTGGGGTCGGTCGCTTGGCTTCCCCTGGCGTCA
GCGACCCAGGGTAACCTCCTCCACTGCTGCGTGCCGTGCAGGCCTGCCTGTGTGAGAGCCACG
TGTGCCGCGCTCTGGGCACAGCCTTGGAAAGTCAGGACCGCGACGGCAGCAGAGCAGAAACCT
TACAGAAACATGAAGCCCTCAACCATCTGCTACTCAGTTATTTCGGGGCTGACGGCGGCTTCTA
GAACATCCAGGTGTTCTGCAGATGCGAGA ACTCATCCTGTAGTCACCAGATGGAGTCCCAAAC
AGCCAAGCAGATGTAAGGCCTGTGCTGTGGCTCTGAGGCCCTGAATACAGAAGGGTCACTTTC
TTAGTGGCCAAAGAGCAGTTGTTGACATTGATGTCTAATTATTGAACACGACCAGTCATTTTA
CTGAGCTGCAGTGAGGAAACACTGACCATAGAAGATCAAGCCAAATGAGGGATTGCAAATTTTC
CTGATTCTTTTGAATTAGGATTCCAGATGGGGGCCCTCATTCTACAGCCCCCAACATTCCTAT
AGCCGTTATCACTGCCATCACCCTGCCACCAGCATCTTCTTGAGATTCCACCCCTGCTCCC
CAGAGACTTCCTGCTTTGAAAGTGAGCAGAAAGGAAGCTCTCAGAAAAATCTCTAGTGGTGGC
TGCCGTCGCTCCAGACAATCGGAATCCTGCCTTACCACCATGGGCTGGCTTTTTCTAAAGGT
TTTGTGTCGGGGAGTGAGTTTCTCAGGATTTCTTTATCCTCTTGTGGATTTTTTGCATCAGTGG
GAAAACAAGAGGACAGAAGCCAAACTTTGTGATTATTTTGGCCGATGACATGGGGTGGGGTGA
CCTGGGAGCAAACCTGGGCAGAAACAAAGGACACTGCCAACCTTGATAAGATGGCTTCGGAGGG
AATGAGGTTTGTGGATTTCATGCAGCTGCCTCCACCTGCTCACCCCTCCCGGGCTTCCTTGCT
CACCGGCCGGCTTGGCCTTCGCAATGGAGTCACACGCAACTTTGCAGTCACCTTCTGTGGGAGG
CCTTCCGCTCAACGAGACCACCTTGGCAGAGGTGCTGCAGCAGGCGGGTTACGTCACCTGGGAT
AATAGGCAAATGGCATCTTGGACACCACGGCTCTTATCACCCCAACTTCCGTGGTTTTGATTA
CTACTTTGGAATCCCATATAGCCATGATATGGGCTGTACTGATACTCCAGGCTACAACCACCC
TCCTTGTCAGCGTGTCACAGGGTGATGGACCATCAAGGAACCTTCAAAGAGACTGTTACAC
TGACGTGGCCCTCCCTCTTTATGAAAACCTCAACATTGTGGAGCAGCCGGTGAACCTTGAGCAG
CCTTGCCCAGAAAGTATGCTGAGAAAGCAACCCAGTTCATCCAGCGTGCAAGCACCAGCGGGAG
GCCCTTCCTGCTCTATGTGGCTCTGGCCACATGCACGTGCCCTTACCTGTGACTCAGCTACC
AGCAGCGCCACGGGGCAGAAGCCTGTATGGTGCAGGGCTCTGGGAGATGGACAGTCTGGTGGG
CCAGATCAAGGACAAAGTTGACCACACAGTGAAGGAAAACACATTCCTCTGGTTTACAGGAGA
CAATGGCCCGTGGGCTCAGAAGTGTGAGCTAGCGGGCAGTGTGGGTCCCTTCACTGGATTTTG
GCAAACCTCGTCAAGGGGGAAGTCCAGCCAAGCAGACGACCTGGGAAGGAGGGCACCGGGTCCC
AGCACTGGCTTACTGGCCTGGCAGAGTTCCAGTTAATGTCACCAGCACTGCCTTGTTAAGCGT
GCTGGACATTTTTTCCAACCTGTGGTAGCCCTGGCCCAGGCCAGCTTACCTCAAGGACGGCGCTT
TGATGGTGTGGACGTCTCCGAGGTGCTCTTTGGCCGGTCACAGCCTGGGCACAGGGTGCTGTT
CCACCCCAACAGCGGGGCAGCTGGAGAGTTTGGAGCCCTGCAGACTGTCCGCCTGGAGCGTTA
CAAGGCCTTCTACATTACCGGTGGAGCCAGGGCGTGTGATGGGAGCATGGTGCCTGAGCTGCA
GCATAAGTTTCTCTGATTTTCAACCTGGAAGACGATACCGCAGAAGCTGTGCCCTAGAAAG
AGGTGGTGCAGGAGTACCAGGCTGTGCTGCCCAGGTCAGAAAGGTTCTTGCAGACGTCTCCA
AGACATTGCCAACGACAACATCTCCAGCGCAGATTACACTCAGGACCCTTCAGTAACTCCCTG
CTGTAATCCCTACCAAATTGCCTGCCGCTGTCAAGCCGCATTAACAGACCAATTTTTATTCCAC
GAGGAGGAGTACCTGGAAATTAGGCAAGTTTGCTTCCAAATTTTCATTTTTACCCTCTTTACAA
ACACACGCTTTAGTTTAGTCTTGGAGTTTAGTTTTGGAGTTAGCCTTGCATATCCCTTCTGTA
TCCTGTCCCCCTCCACGCCGACCCGAGAGCAGCTGAGCTGCGCTGGCTCTGGGCAGGGAGTG
TGCCTTAATGGGAAGCACACGGGCTTTGGAGTCAGGCACAGGTGCCAGCTCCAGCTTTTGAAC
TTGGGCAATTGTTTAACTAACCTGCAAGTTGATTTTGAGGGTTAAATAAAGGCATACATGAA
AATGCCTGGCAACTTTAAAAA

56/550

FIGURE 56

MGWLFLKVLLAGVSFSGFLYPLVDFCISGKTRGQKPNFVII LADDMGWGDLGANWAETKDTAN
LDKMASEGMRFVDFHAAASTCSPSRASLLTGRLGLRNGVTRNFAVTSVGGLPLNETTLAEVLQ
QAGYVTGIIGKWHLGHHGSYHPNFRGFDYYFGIPYSHDMGCTDTPGYNHPPCPACPQGDGPSR
NLQRDCYTDVALPLYENLNIVEQPVNLSLAQKYAEKATQFIQRASTSGRPFLLYVALAHMHV
PLPVTQLPAAPRGRSLYGAGLWEMDSL VGQIKDKVDHTVKENTFLWFTGDNGPWAQKCELAGS
VGPFTGFWQTRQGGSPAKQTTWEGGHRVPALAYWPGRVPVNVSTALLSVLDIFPTVVALAQA
SLPQGRRFDGVDVSEVLFGRSQPGHRVLFHPNSGAAGEFGALQTVRLERYKAFYITGGARACD
GSMVPQLQHKFPLIFNLEDDTAEAVPLERGGAEYQAVLPEVRKVLADVLQDIANDNISSADYT
QDPSVTPCCNPYQIACRCQAA

Important features of the protein:**Signal peptide:**

amino acids 1-16

Transmembrane domain:

amino acids 353-373

N-glycosylation sites.

amino acids 117-120, 215-218, 356-359, 397-500

N-myristoylation sites.

amino acids 12-17, 33-38, 52-57, 97-102, 101-106, 113-118, 158-163, 328-333, 388-393, 418-423, 435-440, 436-441

Amidation site.

amino acids 382-385

Sulfatases signature 2.

amino acids 129-138

57/550

FIGURE 57

TGGACAAGACACCTCCAGGAGCCCAGCTCACAGCCACCGGTACCTTCTTCCAGGACAAGCTGG
GGGCTCCATGGGCGCCTGAGGGCCAGGCGCCAGGGCCGTGGGCACGAGT**ATG**GTGAGACACC
AGCCCCCTGCAGTACTACGAGCCACAGCTGTGCCTCTCCTGCCTCACGGGCATCTACGGCTGCC
GTTGGAAGCGCTACCAGCGCTCCCATGATGATACCACACCGGGCACAGCGCCATTCTGCATG
TGGGGGCTGTGGCAGCAGTCACCATGCTCTCCTGGATCGTGGCAGGACAGTTCCGCCGTGCAG
AGCGGACCTCCTCCCAGGTGACCATTCTCTGTACCTTCTTCACCGTGGTGTTCGCCCTCTACC
TGGCCCCCTCTACCATCTCCTCTCCCTGCATCATGGAGAAGAAAGACCTCGGCCCCAAGCCTG
CTCTCATTGGCCACCGCGGGGCCCCCATGCTGGCTCCAGAGCACACGCTCATGTCTTCCGGA
AGGCCCTCGAGCAGAAGCTGTACGGGCTCCAGGCTGACATTACCATCAGCCTGGACGGCGTGC
CCTTCCTCATGCATGACACCACCCTGCGGCGCACCAACAGTGGAGGAGGAGTTCCCGGAGC
TGGCCCGCAGGCCTGCCTCCATGCTTAAGTGGACCACCCTGCAGAGACTCAACGCTGGCCAGT
GGTTCCTGAAGACTGACCCCTTCTGGACAGCCAGCTCCCTGTCACCCTCCGACCACAGAGAGG
CCCAGAACCAGTCCATCTGCAGCCTGGCAGAGCTCCTGGAGCTGGCCAAGGGCAATGCCACAC
TGCTGCTCAACCTGCGTGACCCGCCCCGGGAGCACCCCTACCGCAGCAGTTTTATCAACGTGA
CTCTGGAGGCGGTGCTGCACTCCGGCTTCCCCAGCACCAGGTCATGTGGCTGCCTAGCAGGC
AGAGGCCCTGGTGCGGAAGGTGGCTCCCGGCTTCCAACAGACATCAGGCTCCAAGGAGGCAG
TCGCCAGCCTGCGGAGAGGGCCACATCCAGCGGCTGAACCTGCGCTACACTCAGGTGTCCCGCC
AGGAGCTCAGGGACTACGCGTCTGGAACCTGAGTGTGAACCTCTACACAGTCAACGCACCGT
GGCTCTTCTCCCTGCTGTGGTGTGCGGGGGTCCCATCCGTACCTCTGACAACTCCCACACCC
TGTCACAGGTGCCTTCCCCCTCTGGATCATGCCCCCGGACGAGTACTGTCTCATGTGGGTCA
CTGCCGACCTGGTCTCCTTCACCCTCATCGTGGGCATCTTCGTGCTCCAGAAGTGGCGCCTGG
GTGGCATAACGGAGCTACAACCCTGAGCAGATCATGCTGAGTGCTGCGGTGCGCCGGACCAGCC
GGGACGTCAGCATCATGAAGGAGAAGCTTATTTTCTCAGAGATCAGCGATGGTGTAGAGGTCT
CCGATGTGCTCTCCGTATGTTTACAGACAACAGTTATGACACATATGCCAACAGCACCGCCACCC
CTGTGGGCCCCCGAGGGGGTGGCAGCCACACCAAGACCCTCATAGAGCGGAGTGGGCGT**TAGC**
TGAAGACATGTCTGTCCCACCTGTACCTGACACAGAAGCTGGGGAGCCTAGGAGAGCTGGTGG
AAGTGTGTCTGAACTCGGAGTGCTCTGGGAGCGGGCTCCACAGCCTCCTTGTGGGCTCCAGCC
CCTTGTACAGCCGCAGCCTCTCTTGAGGGGGACTCCCTGTCTCCTGAGGCCAGCTGGGCCAGG
ACTCCATCCTTTCAGATGCCCTGCAGGCCTGGGGCTCCTTCTGGGAAGTATGGGGCCTAGGG
CTTGGTCCCCCTCTTCTGAGGCCCTCTCCTGTATCCCGACCTGGAAGCTTTGATGGGTCTAGG
GCCATGCCATACCCCTGTGGCAATGGAGTGTGTGGATGCTCACCTGTGCCATCTGTCTCCTCCT
GTCTGTGCCAGGAGGCACCTGAGTTCTCTGCTGTTATCCTGCCCCAAGGGCCTGGGCCGAGCC
TCTACCTGAAGCAACTCTGCTCTTCTGTCTCAGTCTCAAAGCACAAGGAGGTTACGCCAGGAG
GAAGCCAGCTGCAATGTGGAGACACGTCCTCCTCCCCAACCCACCTCATGCCACCGCCAACCC
CCTGCCCCAGGAGCGGGCCTGAGCCACGTCCCCTAGGAGCAGCTGGAGATGGCCAAAAGAGTG
AGCTCAGGACTACTGGATCCCATGCCAGGTGTCCAGCAGACCTCAAGGCAGAAGGGTCACTT
AACCCAGGAGTCCACAGACTGATGTGACCTCAGGTTCCACATCAGTGGCCACAGGGCAGGGC
CCACCTGGTAGAAGTGTCTGGATATGGCCAGGGTGGGTGTGTGGCTAAGTGGGCCTGAACAG
AGGGAACCTAGGGCCCTTGCCCAATGTGATTAAAGCTGCCATCTTGAAA

58/550

FIGURE 58

MVRHQPLQYYEPQLCLSLCTGIYGCRWKRYQRSHDDTTPGTAPFLHVGAVAAVTMLSWIVAGQ
FARAERTSSQVTILCTFFTIVVFALYLAPLTISSPCIMEKKDLGPKPALIGHRGAPMLAPEHTL
MSFRKALEQKLYGLQADITISLDGVFPLMHDTTLRRTTNVEEEFPELARRPASMLNWTTLQRL
NAGQWFLKTDPFWTASSLSPSDHREAQNQSICSLAEELLELAKGNATLLLNLRDPPREHPYRSS
FINVTLEAVLHSGFPQHQMVLPSRQRPLVRKVAPGFQQTSGSKEAVASLRGHIQRLNLRYT
QVSRQELRDYASWNLSVNLYTVNAPWLFSLLWCAGVPSVTSDNSHTLSQVPSPLWIMPPDEYC
LMWVTADLVSFLLIVGIFVLQKWRLGGIRSYNPEQIMLSAAVRRTSRDVSIMKEKLIFFSEISD
GVEVSDVLSVCSDNSYDTYANSTATPVGPRGGGSHTKTLIERSGR

Important features of the protein:**Signal peptide:**

amino acids 1-24

Transmembrane domains:

amino acids 47-61, 77-93, 335-350, 380-399

N-glycosylation sites.

amino acids 182-186, 217-221, 233-237, 255-259, 329-333, 462-466

Tyrosine kinase phosphorylation site.

amino acids 130-139

N-myristoylation sites.

amino acids 21-27, 48-54, 294-300, 404-410, 442-448, 473-479

59/550

FIGURE 59

CCTGAGCAAACACAGCAGCCCGAGTGTTCCTCAAGGCCAAAATGCTGAGAACGTCCACTCCTAA
TCTGTGTGGTGGTCTGCATTGCCGGGCCCCCTGGCTCTCTTCTGGCATTCTCTGCCTCTGCCT
CATATTCTTGTAGGCCAGGTGGGCTTGCTGCAGGGACACCCCCAGTGCCTGGATTACGGGCC
CCCTTTCCAGCCCCCTCTGCACCTTGAGTTTGTCTCTGACTATGAGTCCTTCGGCTGCTGTGA
TCAGCACAAGGACCGCCGATCGCTGCCCCGCTACTGGGACATCATGGAATATTTTGATCTGAA
GAGACATGAGCTGTGTGGAGATTACATTAAAGACATCCTTTGCCAGGAGTGTCTCGCCCTACGC
AGCCCACCTCTACGACGCCGAAAACACCCAGACGCCTCTCCGGAATCTCCCGGGCCTCTGCTC
TGATTACTGCTCTGCCTTCCATTCTAACTGTCACTCAGCCATTTCCCTGCTGACCAATGACCG
CGGCCTCCAGGAGTCTCATGGAAGGGACGGTACCCGCTTCTGCCACCTCCTGGACCTTCCTGA
CAAGGACTATTGCTTCCCTAATGTCTGAGGAACGACTATCTCAACCGCCACCTGGGCATGGT
GGCCCAAGATCCTCAGGGCTGCCTGCAGCTCTGCCTGAGCGAGGTGGCCAACGGGCTGAGGAA
CCCCGTCTCCATGGTCCATGCTGGGGACGGCACCCATCGCTTCTTTGTTGCCGAGCAGGTAGG
AGTGGTGTGGGTCTACCTCCCTGATGGGAGTCGCTGGAGCAACCCTTCCTGGACCTCAAGAA
CATCGTGTTGACCACCCCATGGATCGGGGATGAGAGAGGCTTCTTGGGGTGGCTTTTCACCC
CAAATTCCGCCACAATCGCAAGTTCTATATTTATTATTCGTGCCTGGACAAGAAGAAGGTAGA
AAAGATCCGAATTAGTGAGATGAAGGTTTCTCGGGCTGATCCTAACAAAGCTGACCTGAAATC
AGAGAGGGTCACTTGGAGATTGAAGAACCAGCCTCAAACCATAATGGCGGACAACCTTCTTTT
TGGCCTGGATGGCTATATGTACATATTCACTGGGGACGGGGGACAGGCTGGAGATCCCTTTGG
CCTGTTTGGAAATGCTCAGAACAAAAGTTCCTGCTGGGAAAAGTTTTAAGGATCGATGTGAA
CAGGGCAGGCTCACATGGCAAGCGGTACCGAGTCCCCTCGGACAATCCATTTGTTTCTGAGCC
AGGGGCCCCACCCCGCCATCTATGCCTATGGGATCAGGAACATGTGGCGTTGTGCTGTGGACCG
AGGGGACCCCATCACGCGCCAGGGCCGAGGCGGGATATTCTGTGGGGACGTGGGCCAGAACAG
GTTTGAAGAGGTTGACCTCATTTTGAAGGTGGAACTATGGCTGGAGAGCAAAGGAAGGGTT
TGCATGTTATGACAAAAAACTTTGTACAAATGCCTCTTTGGATGATGTTCTGCCAATCTATGC
TTATGGCCATGCAGTGGGGAAGTCAGTCACTGGAGGTTATGTCTATCGTGGTTGTGAATCCCC
AAATCTCAATGGCCTGTATATCTTTGGAGACTTCATGAGTGGTCGACTTATGGCTTTCAGGA
AGATAGAAAAAACAAGAAATGGAAGAAGCAGGATCTTTGCCTGGGCAGCACACGTCCTGTGC
CTTCCCAGGGCTGATCAGCACCCATAGCAAGTTCATCATCTCCTTTGCTGAAGATGAAGCAGG
GGAGCTGTATTTCTGGCGACCTCTTACCCAAGTGCCTATGCACCACGTGGATCTATTTACAA
GTTTGTGACCCCTCAAGGCGAGCACCCCCAGGCAAGTGCAAATACAAGCCAGTGCCCGTGAG
AACCAAGAGTAAGCGGATCCCGTTCAGACCACTCGCCAAGACAGTCTTGACTTGCTAAAGGA
ACAATCAGAGAAAGCTGCTAGAAAATCTTCCAGTGCAACCTTAGCTTCTGGCCCAGCCCAGGG
TTTGTCTGAGAAAGGCTCCTCCAAGAAGCTGGCTTCTCCTACAAGCAGCAAGAATACATTGCG
AGGGCCTGGTACAAAGAAGAAAGCCAGAGTGGGGCCCCACGTCCGCCAGGGCAAGAGGAGGAA
GAGCCTGAAAAGCCACAGTGGCAGGATGAGGCCATCAGCAGAGCAGAAGCGAGCTGGCAGAAG
TCTCCCTTGACCTATTGGTCAAGGTGGCCGACAGGGTGACGTGAGAGAGGAGAGCCACCTCAT
CAAATGAAAGTCACTGCTGAATAAAGACCTTAGAAGTCTGGGAAGCCAGGGTAGAGGTGGGGC
AGGGCGGTTTTCTCTCCCTGGGAAATCTTGCTGTCTACTGAATAAATAAATGCACCTTCTCT
GTATGCAGTGCTTCTGTGGGAGACCATATCCAGATTGCTGGTGCACCTGGGTTATGGTAAGC
ACTAGTCCATGAGCCTGCTTGGAATCACACTGGATGTCTCCGTTTTGTCTTGTAATGCCTAC
AACCTGAGGTAATAAATCAACATTTGCTCA

60/550

FIGURE 60

MLRTSTPNLCGGLHCRAPWLSSGILCLCLIFLLGQVGLLQGHQPCLDYGPPFQPPLHLEFCSD
YESFGCCDQHKDRRIAARYWDIMEYFDLKRHELCDYIKDILCQECSPYAAHLYDAENTQTPL
RNLPGLCSDYCSAFHSNCHSAISLLTNDRLQESHGRDGTFRCHLLDLDPDKDYCFPNVLRNDY
LNRHLGMVAQDPQGCLQLCLSEVANGLRNPVSMVHAGDGTFRFFVAEQVGVVWVYLPDGSRL
QPFLDLKNIVLTTPWIGDERGFLGLAFHPKFRHNRKFYIYYSCLDKKKVEKIRISEMKVSRAD
PNKADLKSERVILEIEEPASNNHGGQLLFGLDGYMYIFTGDGGQAGDPFGLFGNAQNKSSLLG
KVLRIDVNRAGSHGKRYRVPSDNPFVSEPGAHPAIYAYGIRNMWRCAVDRGDPITRQGRGRI
CGDVGQNRFEEDLILKGGNYGWRAKEGFACYDKKLCHNASLDDVLPYAYGHAVGKSVTGGY
VYRGCESPNLNGLYIFGDFMSGRLMALQEDRKNKKWKKQDLCLGSTTSFAFPGLISTHSEKFI
SFAEDEAGELYFLATSYPSAYAPRGSYKFVDPSRRAPPGKCKYKVPVVRTKSKRIPFRPLAK
TVLDLLKEQSEKAARKSSSATLASGPAQGLSEKSSKKLASPTSSKNTLRGPGTKKKARVGPH
VRQGKRRKSLKSHSGRMRPSAEQKRAGRSLP

Important features of the protein:**Signal peptide:**

amino acids 1-41

Transmembrane domain:

amino acids 17-36

N-glycosylation sites.

amino acids 372-376, 480-484

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 645-649, 699-703

Tyrosine kinase phosphorylation site.

amino acids 81-89

N-myristoylation sites.amino acids 11-17, 37-43, 156-162, 165-171, 357-363, 365-371,
368-374, 408-414, 459-465, 548-554, 557-563**Amidation sites.**

amino acids 391-395, 696-700

Cell attachment sequence.

amino acids 428-431

Leucine zipper pattern.

amino acids 25-47

61/550

FIGURE 61

CTCCATTAAACCACCACCAGCTCCCCAAGCCACCCCTTCAGCCATGAAGTTCCTGCTCCTGGT
CTTGGCAGCCCTCGGATTCCTGACCCAGGTGATCCCAGCCAGTGCAGGTGGGTCAAAATGTGT
GAGTAACACCCCAGGATACTGCAGGACATGTTGCCACTGGGGGGAGACAGCATTGTTTCATGTG
CAACGCTTCCAGAAAATGCTGCATCAGCTACTCCTTCCTGCCGAAGCCTGACCTACCACAGCT
CATCGGTAACCACTGGCAATCAAGGAGAAGAAACACACAAAGGAAAGACAAGAAGCAACAAAC
GACCGTAACATCATAATAACCACTGCTATCGCCTCCACCAACTCAGAGAAATATCATTTCCAC
AGTTCCAATTCTCCTACATTGCTGAGTACTAGCCAAGGCTCCTCTTTATGGGGCAGATATCT
ATAGCCAACCCCAAACCTTCTGTCTTCTATCATTCTGTCAATTCATCTAGTAACTAATTTGGAG
TTTGTATCTATCTTACGAGAACAATCATCATGCAGATTCTGTCCACAGGGGATCTGTCTAGTTG
GGTCTCCAAATGAAAAATGTCAAGACAGAATTGGACATGCAAAAGATTGACTGGGAGAACAC
ACCTCTGATGGACAAAGGTGAGACAGAGCAGCCACAGGCAGGGAGAGCCTTCAGACTGCAACG
CTGGCCTGATACGTGTCAAAGGAGAGAGGGATAGAGGAGGATTGAATAGAAGGAGACTAAGAC
TGCAGCTCTAAGAAAGTCTCAGCCAAACAGATGGGGAGGCCCAAAGCAAGGCTTGCCCCCTCAG
AGGAGCTCACGCAGGGCAGGAATAGCCAGGTTCTCATATCCCAGGGGTTTCAGACTTGGCTGAG
AACAGCCCCTGGGAGAACATGGGGTGACTGCTACCATAGGTCTGGAAGTATGAGGCTGTCCACC
AACTATCCCCTTGAAGCAAGTTCTCTTGAAAGGAAATCTAAACAGTGCACCCCCATGGCTGCC
ACGGAGTATAAGGAGGGAGAGAAAGGAGCTGAAAGTCTAGGTTTGGCCAGCTAGGTAGACTGA
CTTGTGAGGTATTTATTTATTCATTTGAGTAACAAAGCAGACAGAATACATAGCCACCATTGG
TAGTACACCCCAAAGCAAGGATGGCATGATGCTGGTGACTCAAACGTGCCTACTCATGGTGT
CAAATTGGCATAATCCTCTTGGAAGCTGTGTGGAAATAAGCACAGAGAAGCAGAACTCTAAT
TGCTTAATCCACTAAACATTACTTCTGGGAATTGGCTCATCATAAATTATCCAAGAGAAAGCA
CAAAGTTATGGGCACAAAGGTTTTCCATATAATATTATTTAAATGCTGAGAAATGAAAAAA
TCTAAATGGTGAAATATATACTAATGCCATCTATAAATACAAACAAATAGAATGTTTATAGAA
TAATGGAACATAATAACATTATTCAAAATTGCATTTATGCTATAGTTGTCAAATGTCTCCT
TATATGATACAAAACCTCATGAAAATTATGACTTTTTTGTGGTTGGAAAGCAGAATTATGCA
TAAATTTCTCTTACAGTTCGATGCCCATTAGTTTTATATAACATTTATTTGACACGTAAGTGA
CTTCTATCTGAGAAGAACAACCAAAACACTCAGGCCTAAATAATTAACCGGTCTTAAAAA
CTAGCAAACCAGATAAGAAAAGATGTTAATGCCCATTCCTTAACCTTATGTCTTAGACCAAAT
TAATTCATAGATGGTTTTTAAATGACAGTGTAAGTAAAGTATTAAAGATTTGTGTGGTCAA
TATTCAATTTAAGAGCAAGGAAATTCTTATAAATATAACAATAGAGGCAGAACTCATGTAAGA
ATAAATTGATTAGGTGGTATTAAATATTAAGTTCTTATGTATGTCAAAGATATCATTTTGAA
ATTCATCCATCTTATTGGGTATTGCAGGAGTTCATTCCTTTTTTGTATATAAATACTCTCCGT
CATATGAATAGTATTCATTTGTATACTGGTTTGTGATGGACATTTGGGTGTTCCAGTTTA
TGGCTATTACAAATAAAGCTTCTATGAACATTTATGTACA

62/550

FIGURE 62

MKPELLLVLAALGFLTQVIPASAGGSKCVSNTPGYCRTCCHWGETALFMCNASRKCCISYSFLP
KPDLPQLIGNHWQSRRRNTQRKDKKQQTTVTS

Important features of the protein:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 1-22

N-glycosylation site.

amino acids 50-53

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 79-82

N-myristoylation site.

amino acids 23-28

63/550

FIGURE 63

GCGGAGCGCCTGGGAGAGGAGAAGGAGCCGACCTGCCGAGATGGAGGCGACCGGCACCTGGGC
GCTGCTGCTGGCGCTGGCGCTGCTCCTGCTGCTGACGCTGGCGCTGTCCGGGACCAGGGCCCCG
AGGCCACCTGCCCCCGGGCCACGCCGCTACCACTGCTGGGAAACCTCCTGCAGCTACGGCC
CGGGGCGCTGTATTACAGGGCTCATGCGGCTGAGTAAGAAGTACGGACCGGTGTTACCATCTA
CCTGGGACCCTGGCGGCCTGTGGTGGTCCTGGTTGGGCAGGAGGCTGTGCGGGAGGCCCTGGG
AGGTCAGGCTGAGGAGTTCAGCGGCCGGGGAACCGTAGCGATGCTGGAAGGGACTTTTGATGG
CCATGGGGTTTTCTTCTCCAACGGGGAGCGGTGGAGGCAGCTGAGGAAGTTTACCATGCTTGC
TCTGCGGGACCTGGGCATGGGGAAGCGAGAAGGCGAGGAGCTGATCCAGGCGGAGGCCCGGTG
TCTGGTGGAGACATTCCAGGGGACAGAAGGACGCCCATTCGATCCCTCCCTGCTGCTGGCCCA
GGCCACCTCCAACGTAGTCTGCTCCCTCCTCTTTGGCCTCCGCTTCTCCTATGAGGATAAGGA
GTTCCAGGCCGTGGTCCGGGCAGCTGGTGGTACCCTGCTGGGAGTCAGCTCCCAGGGGGGTCA
GACCTACGAGATGTTCTCCTGGTTCCCTGCGGCCCTGCCAGGCCCCCAACAAGCAGCTCCTCCA
CCACGTCAGCACCTTGCGCTGCCTTCACAGTCCGGCAGGTGCAGCAGCACCAGGGGAACCTGGA
TGCTTCGGGCCCCGCACGTGACCTTGTGATGCCTTCCTGCTGAAGATGGCACAGGAGGAACA
AAACCCAGGCACAGAATTCACCAACAAGAACATGCTGATGACAGTCATTTATTTGCTGTTTGC
TGGGACGATGACGGTCAGCACCACGGTCGGCTATACCCTCCTGCTCCTGATGAAATACCCTCA
TGTCCAAAAGTGGGTACGTGAGGAGCTGAATCGGGAGCTGGGGGTGGCCAGGCACCAAGCCT
AGGGGACCGTACCCGCTCCCTTACACCGACGCGGTTCTGCATGAGGCGCAGCGGCTGCTGGC
GCTGGTGCCCATGGGAATACCCCGCACCCCTCATGCGGACCACCCGCTTCCGAGGGTACACCCT
GCCCCAGGGCACGGAGGTCTTCCCCCTCCTTGGCTCCATCCTGCATGACCCCAACATCTTCAA
GCACCCAGAAGAGTTCAACCCAGACCGTTTCCTGGATGCAGATGGACGGTTCAGGAAGCATGA
GGCGTTCCCTGCCCTTCTCCTTAGGGAAGCGTGTCTGCCTTGGAGAGGGGCTGGCAAAGCGGA
GCTCTTCCCTCTTCTTACACCACCATCCTACAAGCCTTCTCCCTGGAGAGCCCGTGGCCGCCGGA
CACCTGAGCCTCAAGCCCACCGTCAGTGGCCTTTTCAACATTCCCCCAGCCTTCCAGCTGCA
AGTCCGTCCCACCTGACCTTCACTCCACCACGCAGACCAGATGAAGGAAGGCAACTTGGAAAGTG
GTGGGTGCCCAGGACGGTGCCTCCAGCCTCAACAGTGGGCATGGACAGGGTTAATGTCTCCAG
AGTGTACACTGCAGGCAGCCACATTTACACGCCTGCAGTTGTTTTCCGGAGTCTGTCCACGG
CCCACACGCTCACTTGACTCATGCTGCTAAGATGCACAACCGCACACCCATACACAACCTACAA
GGGCCACAAAGCAACTGCTGGGTTAGCTTTCCACAGACATAAATATAGTCCATCTGCAATCAC
AAGCACATAGCCAGGTAACCCACCAACTCCCCTGGATCTGCAGCCACACGTGGGAGTCTGGC
TGTCACCTTCACAAGCCACAGAAACGGCCACACATGTTACAGCTCACACGCCCTCTCCATT
ATCGAACTTCTCAGTGTCCCTGTCCCTGGTGCCTGGCACAGGGAACAGCATGCCCCCTCCGGG
GTCATGCCACCCAGAGACTGTCGCTGTCTATGGCCCCAACTCATGCTCCCTCTCTTGGCTACA
CCACTCTCCCAGCCTGTGACCACCGATGTCCACACACCCCCAACCACTTGTCACACAGCTAC
CCACGTACAACATCGTCCTGGCTCCCCAGAGTATCTTCCCCTGAGACACGCCGCCCCACAG
AGGCACAGTCCCCAGCCACCTCTGCAACTGCAGCCCTCAGTCACCCCTTTTTAAGCACCCCTGA
TTCTACCAAATGCAAACACATCTGGGTCTGCGATTATGCACAGAGACTTTGGACATACGAGGA
CCCTCAGACCCGAGGAACACCTGCCCCAACCCCAACACGTGCTTATGTAACCACGTGGAAAGCG
GCCCTGCTGCCCTCCACACACACATACACACTCACTGATCTACAGCCCCTGTTGGCGTCA
GAGTCCCCACTAGACCCAGTGGAAAGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAATTCACC
CTGTCAGGGAGTGAGCCGATCTGACGTTCTTGTGACTTAAGGGTCCGGCTTGGGAATTAAA
GTTTGTTTCTGGCCTTTAGCCTAAAAA

64/550

FIGURE 64

MEATGTWALLLLALALLLLTLALSGTRARGHLPPGPTPLPPLLGNLLQLRPGALYSGLMRLSKK
YGPVFTIYLGPRPVVVLVGQEAVERALGGQAEFFSGRGTVMLEGTFDGHGVFFSNGERWRQ
LRKFTMLALRDLGMGKREGEELIQAEARCLVETFQGTTEGRPFDPDLLAQATSNVVCSSLFGL
RFSYEDKEFQAVVRAAGGTLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHHVSTLAAFTVRQV
QQHQGNLDASGPARDLVDAFLLKMAQEEQNPGTEFTNKNMLMTVIYLLFAGTMTVSTTVGYTL
LLMKYPHVQKWVREELNRELGAGQAPSLGDRTRLPYTDAVLHEAQRLALVPMGIPRTLMRT
TRFRGYTLPQGTVEFPLLGSILHDPNIFKHPEEFNPDREFLDADGRFRKHEAFLPFSLGKRVCL
GGLAKAELFLFTTILQAFSLESPPDLSLKPTVSGLFNIPPAFQLQVRPTDLHSTTQTR

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 294-313

Glycosaminoglycan attachment site.

amino acids 99-103

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 128-132

N-myristoylation sites.

amino acids 51-57, 109-115, 115-121, 188-194, 207-213, 257-263,
284-290, 339-345, 370-376, 444-450

Amidation sites.

amino acids 140-144, 435-439

Leucine zipper pattern.

amino acids 32-54, 39-61

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 433-443

65/550

FIGURE 65

CGGACGCGTGGGGCCGT**ATG**CGCGGCTCTGTGGAGTGCACCTGGGGTTGGGGGCACTGTGCCC
CCAGCCCCCTGCTCCTTTGGACTCTACTTCTGTTTGCAGCCCCATTTGGCCTGCTGGGGGAGA
AGACCCGCCAGGTGTCTCTGGAGGTCATCCCTAACTGGCTGGGCCCCCTGCAGAACCTGCTTC
ATATACGGGCAGTGGGCACCAATTCCACACTGCACTATGTGTGGAGCAGCCTGGGGCCTCTGG
CAGTGGTAATGGTGGCCACCAACACCCCCACAGCACCCCTGAGCATCAACTGGAGCCTCCTGC
TATCCCCTGAGCCCGATGGGGGCCTGATGGTGCTCCCTAAGGACAGCATTCAGTTTTCTTCTG
CCCTTGTTTTTACCAGGCTGCTTGAGTTTGACAGCACCAACGTGTCCGATACGGCAGCAAAGC
CTTTGGGAAGACCATATCCTCCATACTCCTTGGCCGATTTCTCTTGGAACAACATCACTGATT
CATTGGATCCTGCCACCCTGAGTGCCACATTTCAAGGCCACCCCATGAACGACCCTACCAGGA
CTTTTGCCAATGGCAGCCTGGCCTTCAGGGTCCAGGCCTTTTCCAGGTCCAGCCGACCAGCCC
AACCCCTCGCCTCCTGCACACAGCAGACACCTGTCAGCTAGAGGTGGCCCTGATTGGAGCCT
CTCCCCGGGGAAACCGTTCCCTGTTTGGGCTGGAGGTAGCCACATTTGGGCCAGGGCCCTGACT
GCCCCCTCAATGCAGGAGCAGCACTCCATCGACGATGAATATGCACCGGCCGTCTTCCAGTTGG
ACCAGCTACTGTGGGGCTCCCTCCCATCAGGCTTTGCACAGTGGCGACCAGTGGCTTACTCCC
AGAAGCCGGGGGGCCGAGAATCAGCCCTGCCCTGCCAAGCTTCCCCTCTTCATCCTGCCTTAG
CATACTCTCTTCCCAGTCACCCATFGTCCGAGCCTTCTTTGGGTCCCAGAATAACTTCTGTG
CCTTCAATCTGACGTTCCGGGGCTTCCACAGGCCCTGGCTATTGGGACCAACACTACCTCAGCT
GGTCGATGCTCCTGGGTGTGGGCTTCCCTCCAGTGGACGGCTTGTCCCCACTAGTCCTGGGCA
TCATGGCAGTGGCCCTGGGTGCCCCAGGGCTCATGCTGCTAGGGGGCGGCTTGGTTCTGCTGC
TGCACCACAAGAAGTACTCAGAGTACCAGTCCATAAAT**TAA**GGCCCGCTCTCTGGAGGGAAGG
ACATTACTGAACCTGTCTTGCTGTGCCTCGAACTCTGGAGGTTGGAGCATCAAGTTCCAGCC
GGCCCCCTTCACTCCCCCATCTTGCTTTTCTGTGGAACCTCAGAGGCCAGCCTCGACTTCCTGG
AGACCCCCAGGTGGGGCTTCCTTCATACTTTGTTGGGGGACTTTGGAGGCGGGCAGGGGACAG
GGCTATTGATAAGGTCCCCTTGGTGTTGCCCTTCTTGCACTCTCCACACATTTCCCTTGGATGGG
ACTTGCAGGCCTAAATGAGAGGCATTCTGACTGGTTGGCTGCCCTGGAAGGCAAGAAAATAGA
TTTATTTTTTTTCACAGGGAAAAA

66/550

FIGURE 66

MRGSVECTWGWGHCAFSPLLLWTLLEFAAPFGLLGKTRQVSLEVIPNWLGPQLNLLHIRAVG
TNSTLHYVWSSLGPLAVVMVATNTPHSTLSINWSLLLSPEPDGGLMVLPKDSIQFSSALVFTR
LLEFDSTNVSDTAAPLGRPYPPYSLADFSWNNITDSLDPATLSATFQGHMNDPTRTFANGS
LAFRVQAFSRSSRPAQPPRLLHTADTCQLEVALIGASPRGNRSLEGLEVATLGQGPDCPSMQE
QHSIDDEYAPAVFQLDQLLWGSLSGFAQWRPVAYSQKPGGRESALPCQASPLHPALAYSLPQ
SPIVRAFFGSQNNFCAFNLTFGASTGPGYWDQHYLSWSMLLGVGFPPVDGLSPLVLGIMAVAL
GAPGLMLLGGGLVLLLHHKKYSEYQSIN

N-glycosylation sites:

amino acids 65-69, 95-99, 134-138, 159-163, 187-191, 230-234,
333-337

cAMP- and cGMP-dependent protein kinase phosphorylation site:

amino acids 397-401

Casein kinase II phosphorylation sites:

amino acids 151-155, 249-253, 255-259

N-myristoylation sites:

amino acids 3-9, 63-69, 235-241, 273-279, 292-298, 324-330

Leucine zipper pattern.

amino acids 371-393

67/550

FIGURE 67

CGGGACAGGCGCGTGAGGCCACAACACATGCGTGTATCTTGCTGGGCTATCTTCCCTGCTCTGCCACGCCGGGT
 CTGGAGAAGGGGTTTCAGCCCCAGGACATTTACTGAGAGTCGGCGAATATTGGGAGCCGCGATGTTCCCCCTTCG
 GGCCCTGTGGTTGGTCTGGGCGCTTCTAGGAGTGGCCGGATCATGCCCGAGCCGTGCGCTGCGTGGACAAGTA
 CGCTCACCAGTTCGCGGACTGCGCTTACAAAGAGTTGCGTGAGGTGCCGGAAGGACTGCCTGCCAACGTGACGAC
 GCTTAGTCTGTCCGGAACAAGATCACTGTGCTGCGGCGCGGGGCTTCGCCGACGTACACAGGTACGTCGCT
 GTGGCTGGCGCACAAATGAGGTGCCACCCGTGGAGCCAGGCGCACTGGCCGTGCTGAGTCAGCTCAAGAACCTCGA
 TCTGAGCCACAACCTTCATATCCAGCTTTCCGTGGAGCGACCTGCGCAACCTGAGCGCGCTGCAGCTGCTCAAAAT
 GAACCACAACCGCTGGGCTCTCTGCCCGGGACGCACTCGGTGCGCTACCCGACCTGCGTTCCCTGCGCATCAA
 CAACAACCGGCTGCGTACGCTGGCGCTGGCACCTTCGACCGCTTAGCGCGCTGTACACTTGCAACTCTATCA
 CAATCCCTTCCACTGCGGCTGGCGCTTGTGTGGCTGCGGCTGAGGGGTGCCGGTGTACCGCTGCCCGCTGCGCTG
 GCGGACTCCATTGCTTGTGCTGCGCTCCGCGCTGCGAGGGGTGCCGGTGTACCGCTGCCCGCTGCGCTG
 TGCACCGCCAGCGTGCATCTGAGTGCCGAGCCACCGCTTGAAGCACCGGACCCCACTGCGCGCAGGACTGGC
 GTTCGTGTACACTGCATCGCGACGGCCACCCCTACGCTGCGCTGCAATGGCAACTTCAGATCCCCGGTGGCAC
 CGTAGTCTTAGAGCCACCGGTTCTGAGCGGGAGGACGACGGGTTGGGGCGGAGGAAGGAGAGGGAGAAGGAGA
 TGGGGATTGCTGACGCAGACCCAAAGCCAAACGCGGACTCCAGCACCGCTTGGCCGGCGCCCCAGCCACACC
 GCGCTTCCTGGCCCTCGCAAATGGCTCCCTGTTGGTGGCCCTCCTGAGTGCCAAGGAGGCGGGCGTCTACACTG
 CCGTGACACAATGAGCTGGGCGCAACTCTACGTCAATACGCGTGGCGGTGGCAGCAACCGGGCCCCCAAAACA
 CGCGCTGGCGCGGGGGAGAACCCGACGAGGCCCCGACCTCTGAGCGCAAGTCCACAGCCAAAGGGCCGGGG
 CAACAGCGCTCCTGCGCTTCCAAACCCGAGGGCAAAATCAAAGGCCAAGGCTGGCCAAGGTGAGTCTCGGGGA
 GACCGAGACGGAGCCGGAGGAGGACACAAGTGAGGGAGAGGAGGCCGAAGACCAGATCTCGCGGACCGGCGGA
 GGAGCAGCGCTGTGGCAACGGGGACCCCTCTCGGTACGTTTCTAACACGCGTTCAACCAGAGCGCAGAGCTCAA
 GCCGACGCTCTCGAGCTGGCGCTCATCGCGTGGATGTGGCGGAGCGCGAGGCGCGGGTGCAGCTGACTCCGCT
 GGCTGCGCGCTGGGGCCCTGGGCGCGCGGGCTGGCGGAGCCCGCGACCCGGGCGGCGACCCCTGCGCGCTACT
 CTATCTGTGTCCAGCGGGGGCGGCGCGGCACTGCACTGGTCCCGCTAGAGGAAGGCGTCAACGCTACTGCTT
 CCGCGGCTGCGGCGGGTACCAACTACTCCGTGCTGCTGGCGCTGGCGGGCGAAGCCTGCCACGTGCAAGTGGT
 GTTTTCCACCAAGAAGGAGCTCCCATCGCTGCTGGTCAATAGTGGCAGTGAGCGTATTCTCTCTGGTGTGCGGC
 AGTGCCCTTCTGGGCGCGCGCTGCTGCGCTGCTGGCTAAACACCGGGCAAGCCCTACCGTCTGATCCTGCG
 GCCTCAGGCCCCCTGACCCCTATGGAGAAGCGCATCGCGCAGACTTCGACCCGCGTGTCTGTAACCTCGAGTCCGA
 GAAAAGCTACCCGCGAGGCGGCGAGGCGGGCGCGAGGAGCCAGAGGAGCTGCAGGGGGAGGGCCTTGATGAAGA
 CGCGGAGCAGGGAGACCCAAAGTGGGGACCTGCAGAGAGAGGAGAGCCTGGCGGCTGCTCACTGGTGGAGTCCCA
 GTCCAAGGCCAACCAAGAGGAGTTCGAGGCGGGCTCTGAGTACAGCGATCGGCTGCCCTGGGCGCGGAGGCGGT
 CAACATCGCCAGGAGATTAAATGGCAACTACAGGCGAGCGGCGAGGCTGAACCTCCGCGCGTCCGGCCCCGCCATT
 CCGGACCTCCACCTAGGGTGCTGGGAGCAGCAGTCTAGGGCTGGCAGGACTTATGTCCCCCGTCCCCAACCTTC
 ACCTACTCCTCCCCCTTACTACTCCCCAACCTTGACTACCCAGGAGTCTTATTAGGGAGTGGGCGGATTTACCA
 STCCCTGCTACCCACGGCTGCCATTCCTCTGCGGGCTGAATCCCTTCCCCGCCAAGCACAGTGTATTATCTTAC
 CCCATCGAAGACTCCACCCGAGACGGTGGCGGATATCTATGTCCCTCCATTCCCGTCCGCTGATCTGCGGAAAT
 CCACCCGCGAGCCCCACCGTGGGCTCTGGAGCCAGAGGAAACGAGCGAAGACTTTGGAACCTCGCGGTAA
 CCGGCTGGTTTCGGGGGCCAGCCAAAGGCCAGTGGAGTGTGTGGGGTCCACCTCGACCCCTCTCTCTCTCTTTC
 TTTCTTTCTTTTCTTTTATTTTAAATTTTATTTATTTATTTATTTATTTTACGGAGTCTTGGTCTGTGCG
 CAGGCTGGAGTGCAGTGGCGGATCTCGGCTCACTGCATCTTCCGCTCCCGGTTCAAGCGATTCTCCTGCTC
 AGCCTGCCTAGTAGCTGGGACTACAGGCGCGCGCCACCACGACCAGCTAATTTCTTCTATTTTATAGAGACGG
 GGTTCACCATGTTGGCCAGGATGGTCTGGATCTCTTGACCTCAGGTGATCCATCTGCCTCGGCTCTCAAAGTG
 CTGGGATTACAGGCGTGAGGCAACCSCGCCCCGCCCTCTCCCTTTCAATCCCTACTCCAGAAAGCCGGGATTG
 TGGCAACCCCTAGTTTTTACTTCCAAAGCCTCCTGCGGCGAGGGAACCAATCCTTCTGCTCCACCCCCACC
 CCACCTTCTGGCCAGTTGGAGTCCAGGCCGCTGGCTGGGGCGCTTTCAGCTCCGCGCTCAGATTTTCTGTCTTC
 GTTGTCTTCAAAGACAGCGACATTCGGGCTGCTGCTTAACACCCCTTCCAGCCTCTGGGAAAATCGAGTGTG
 TGTGTGGGGGGTAGGCGAGGAATGCGTTTCTGTGCTCTCTCTCTAACTTAAGCGCGCAGGACCGCGCGCC
 CCTTGGCGGCTGAGCTGTGACTTGGTGGCGGGCAATTTCTGTTGTGCTGTGTTGGGCTTTCGGAGGTCTGT
 GCGCCCAACAGCGCGGCTCCCGCGGCTCCACCCGACCCAGACCTAGCTGGAAAGCGCGGAGGCGGAGGAAGCT
 GACTGTGGCTTCCCGGGCGCGGCTCTCTGGAGGGCTCGCGCCCTAGTTGCGACAAAGCCTGCTCGTGAAGTGTG
 GACTGTGCGACGGGATCCGGATGAGCGCGAGCCCTCCGTCTCGCTCTCGGTCTCGGCTCGCCCCGCCCCAC
 CCGCCCCCTGCTTCGGCGGGAAATCGTGTGTTGCCCGGCGTGTAGTCCCTGACAAGCGTGCCCTGTAGGAGAAAAGTC
 TGTGCTCCTGTGAAGTGTGACCGTGTAGTGTAGGGGGCGGGCGGGGGCGGATGGGCGGGAGGAGGAAGGG
 GAGGGCGCGCGCGGCGGCTCGGGGCGGGGTTCTTTTTCATTTTGAAGAAAGCGTCGGGGTTGGGGTGGGG
 GGAGTTTTCAGTCTCGGGATCAGCCCTCTCCGCGAAGCGCAGCACAAAGCGGGGCTGGGACGGAGTAGCCCCC
 GGAGCCCGTGCCCTTTCTAAACGCGTCTGTATGCAGTCAATAAACAATCGATTTGAAA

68/550

FIGURE 68

MFPLRALWLWVALLGVAGSCPEPCACVDKYAHQFADCAYPEGLPANVTTLSSLSANKI
 TVLRRGAFADVDTQVTSWLAHNEVRTVEPGALAVLSQLKNLDLSHNFISSFPWSDLRNLSALQ
 LLKMNNHNLGSLPRDALGALPDLRLRINNRLRRTLAPGTFDALSALSHLQLYHNPFHCGCGL
 VWLQAWAASTRVSLPEPDSIACASPPALQGVVPVYRLPALPCAPPSVHLSAEPPELEAPGTPLRA
 GLAFVLHCIADGHPTPRLQWQLQIPGGTVVLEPPVLSGEDDGVGAEEGEGEGDGDLLTQTQAO
 TPTPAPAWPAPPATPRFLALANGSLLVPLLSAKEAGVYTCTRAHNELGANSTSIKVAVAATGPP
 KHAPGAGGEPDGQAPTSEKSTAKGRGNSVLPSKPEGKIKGQGLAKVSILGETETETEPEEDTSE
 GEEAEDQILADPAEEQRCGNGDPSRYVSNHAFNQSAELKPHVFELGVIALDVAEREARVQLTP
 LAARWGPGPGGAGGAPRPGRRPLRLLYLCPAGGGAAVQWSRVEEGVNAYWFRGLRPGTNYSVC
 LALAGEACHVQVVFSTKKELPSLLVIVAVSVFLLVLATVPLLGAACCHLLAKHPGKPYRLILR
 PQAPDPMEKRIAADFDPRASYLESEKSYFAGGEAGGEEPEDVQGEGLDEDAEQGDPSGDLQRE
 ESLAACSLVESQSKANQEEFEAGSEYSDRLPLGAEAVNIAQEINGNYRQTAG

Important features of the protein:**Signal peptide:**

amino acids 1-19

Transmembrane domain:

amino acids 587-610

N-glycosylation sites.

amino acids 52-55, 121-124, 337-340, 364-367, 474-477, 563-566

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 397-400

Casein kinase II phosphorylation sites.amino acids 19-23, 202-205, 289-292, 246-249, 411-414, 431-434,
433-436, 440-443, 544-547, 583-586, 650-653, 700-703**N-myristoylation sites.**amino acids 15-20, 48-53, 165-170, 296-301, 351-356, 362-367,
390-395, 419-424, 514-519, 536-541, 557-562, 561-566, 610-615,
661-666, 716-721**Amidation site.**

amino acids 522-525

Prokaryotic membrane lipoprotein lipid attachment sites.

amino acids 10-20, 603-613

69/550

FIGURE 69

GGCGGCGGGAGCAGCGAAGGGGGCGGCAGGGATCCTCCAGGCTGCCGGCTGGGAAGGCCTGGG
CGACCCGGTGTGTGGCGCGCCAGAGCCCCGCGTTTCAGCCCTAGGGAAGGAAGCCAGTTGAG
GGAAGTTCTCCATGAATGTACGTCACA**ATG**ATGATGACCGACCAAATCCCTCTGGAACTGCCA
CCATTGCTGAACGGAGAGGTAGCCATGATGCCCCACTTGGTGAATGGAGATGCAGCTCAGCAT
GTTATTCTCGTTCAAGTTAATCCAGGTGAGACTTTCACAATAAGAGCAGAGGATGGAACACTT
CAGTGCATTCAAGGACCTGCTGAAGTTCCCATGATGTCACCCAATGGATCCATTCCCTCCCAT
CATGTGCCTCCAGGTTATATCTCACAGGTGATTGAAGATAGTACTGGAGTCCGCCGGGTGGTG
GTCACACCCCAGTCTCCTGAGTGTTATCCCCAAGCTACCCCTCAGCCATGTCTCCAACCCAT
CATCTCCCTCCCTATCTGACTCACCATCCACATTTTATTTCATAACTCACACACGGCTTACTAC
CCACCTGTTACCGGACCTGGAGATATGCCGCCTCAGTTTTTTCCCCAGCATCATCTTCCCCAC
ACAATATATGGTGAGCAAGAAATTATACCATTTTATGGAATGTCAAGCTACATCACCCGAGAA
GACCAGTACAGCAAGCCTCCGCACAAAAAACTGAAAGACCGCCAGATCGATCGCCAGAACCGC
CTCAACAGCCCTCCTTCTTCTATCTACAAAAGCAGCTGCACAACAGTATACAATGGCTATGGG
AAGGGCCATAGTGGTGGAAAGTGGCGGAGGCGGCAGCGGTAGTGGTCCCGGAATTAAGAAAACA
GAGCGACGAGCAAGAAGCAGCCCAAAGTCGAATGATTGAGACTTGCAAGAATATGAGTTGGAA
GTAAAGAGGGGTGCAAGACATTCTTTCGGGAATAGAGAAACCACAGGTTTCTAATATTGAGGCA
AGAGCAGTTGTGTTGTCCTGGGCTCCCCCTGTTGGACTTTCCTGTGGACCCACAGTGGTCTT
TCCTTCCCCTACAGTTACGAGGTGGCCTTATCAGACAAAGGACGAGATGGAAAATACAAGATA
ATTTACAGTGGAGAAGAATTAGAATGTAACTGAAAGATCTTAGACCAGCAACAGATTATCAT
GTGAGGGTGTATGCCATGTACAATTCGGTAAAGGGATCCTGCTCCGAGCCTGTTAGCTTCACC
ACCCACAGCTGTGCACCCGAGTGTCTTTCCCCCTAAGCTGGCACATAGGAGCAAAAGTTCA
CTAACCCTGCAGTGGAAAGGCACCAATTGACAACGGTTCAAAAATCACCAACTACCTTTTAGAG
TGGGATGAGGGAAAAAGAAATAGTGGTTTCAGACAGTGCTTCTTCGGGAGCCAGAAGCACTGC
AAGTTGACAAAGCTTTGTCCGGCAATGGGGTACACATTCAGGCTGGCCGCTCGAAACGACATT
GGCACCAGTGGTTATAGCCAAGAGGTGGTGTGCTACACATTAGGAAATATCCCTCAGATGCCT
TCTGCACCAAGGCTGGTTCGAGCTGGCATCACATGGGTACGTTGCAGTGGAGTAAGCCAGAA
GGCTGTTACCCGAGGAAGTGATCACCTACACCTTGAAATTCAGGAGGATGAAAATGATAAC
CTTTTCCACCCAAAATACACTGGAGAGGATTTAACCTGTACTGTGAAAAATCTCAAAAGAAGC
ACACAGTATAAATTCAGGCTGACTGCTTCTAATACGGAAGGAAAAAGCTGTCCAAGCGAAGTT
CTTGTTTGTACGACGAGTCCTGACAGGCCTGGACCTCCTACCAGACCGCTTGTCAAAGGCCCA
GTTACATCTCATGGCTTTAGTGTCAAATGGGATCCCCCTAAGGACAATGGTGGTTCAGAAATC
CTCAAGTACTTGCTAGAGATTACTGATGGAAATTCGAAGGTGAAGTTTTTGGCAATTGTTTT
ATTCAAATCCA**ATAG**CAAGCTCTGTTTTCTAATATAGTAAATGTCTTTATAGTAATAGTGAGT
AATCATTAATTCTAAAGATAGAATTATTATTACAATAAACAACTTTAGTCACATATTGGCAG
TTTTTCTATTTCAAACACAGCACCAGAGATCAGAGTCTACTTGAACTTACATTTGTGTTATT
TAACAATTTTTCTGTATCTTTTTTCATTGGTGTGTTTTGTTTTGTTTATCTTTGTTTTTGTCT
TTGGTTTTGGTTTTGTTTTGTTTTGTTTTGAGATACGATCTCTGTCCACAGGCTGGAGGGC
AGTGGCACAGACATGGCCCATTCAGTCTCAGACTCCTGGGCTTAAGTGACTCTTCTGCCACA
GAAGATGAGGAAGAATACATTTTTTCATAGTGATGGGGTCTCACTATGTTATCTAGGCTGGTCT
CAAACCTCCTGGCCTCAAGCAACCCTCCACCTTGGCCTCCCAAAGTGCTGGGACTATAGACATG
AATCACCACTCAGCTTCCATGTCTTTTTATGAACTAGGGTTCCTAATTAATCAGATAAATT
TGGTATTTTTCATCTCCTAACTTGCCATATGTTTTCTGGAAATTCCTATAAGCAGCCGAGAGTG
GTGGCTCACGCTGTAGTCCCAGCACTTTGGGAGGCTGAGGTGGGTGGTCAGGAGATCAAGACC
ATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGTGTGGTG
GCAGGCACCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGAAGAATTGCTTGAACCCAGCAG
CGGAGGTTGCAGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGTGACAGAGTGAGACTC
TGCTCTCAAAAAAAAAA

70/550

FIGURE 70

MMMTDQIPLELPPLLNGEVAMMPHLVNGDAAQHVLVQVNPGETFTIRAEDGTLQCIQGPAEV
 PMMSPNGSIPPIHVPPGYISQVIEDSTGVRRVVVTPQSPECYPPSYPSAMSPTHHLPPYLTHH
 PHFIHNSHTAYYPPVTGPGDMPPQFFPQHHLPHTIYGEQEIPFYGMSSYITREDQYSKPPHK
 KLKDRQIDRQNRNLNSPPSSIYKSSCTTVYNGYKGKHSGGSGGGSGSGPGIKKTERRARSSPK
 SNDSDLQEYELEVKRVQDILSGIEKPQVSNIQARAVVLSWAPPVGLSCGPHSGLSFPYSYEVA
 LSDKGRDGKYKIIYSGEELCNLKDRLPATDYHVRVYAMYNVKGSCSEPVSFTHSCAPECP
 FFPKLAHRSSSLTLQWKAPIDNGSKITNYLLEWDEGKRNSGFRQCFFGSQKHCKLTKLCPAM
 GYTFRLAARNDIGTSGYSQEVVCYTLGNIPQMPSAPRLVRAGITWVTLQWSKPEGCSPEEVIT
 YTLEIQEDENDNLFHPKYTGEDLTCTVKNLKRSTQYKFRLTASNTEGKSCPSEVLVCTTSPDR
 PGPPTRPLVKGPVTSHGFSVKWDPPKDNNGGSEILKYLLEITDGNSEGEVFGNCFIQIQ

Important features of the protein:**N-glycosylation sites.**

amino acids 69-73, 254-258, 401-405

Glycosaminoglycan attachment sites.

amino acids 229-233, 234-238, 236-240

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 416-420, 535-539

Tyrosine kinase phosphorylation site.

amino acids 319-326

N-myristoylation sites.amino acids 52-58, 227-233, 228-234, 230-236, 231-237, 232-238,
235-241, 239-245, 402-408, 610-616**Amidation site.**

amino acids 414-418

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 290-301

ATP/GTP-binding site motif A (P-loop).

amino acids 546-554

CUB domain proteins profile.

amino acids 294-301

71/550

FIGURE 71

AAGTCATTCAAGTGGATGTGATCTTGGCTCACAGGGGACGATGTCAAGCTCTTCCTGGCTCCTTCTCAGCCTTGTT
GCTGTAACTGCTGCTCAGTCCACCATTGAGGAACAGGCCAAGACATTTTTGGACAAGTTTAACCACGAAGCCGAA
GACCTGTTCTATCAAAGTTCAGTCTGCTTCTTGGGAATTATAACACCAATATTACTGAAGAGAATGTCCAAAACATG
AATAATGCTGGGGACAAATGGTCTGCCTTTTTAAAGGAACAGTCCACACTTGCCCAAATGTATCCACTACAAGAA
ATTGAGAATCTCACAGTCAAGCTTCAGCTGCAGGCTCTTCAGCAAAATGGGTCTTCAGTGCTCTCAGAAGACAAG
AGCAAACGGTTGAACACAATTTCTAAATACAATGAGCACCATCTACAGTACTGGAAAAGTTTGTAAACCCAGATAAT
CCACAAGAATGCTTATTACTTGAACACAGGTTTGAATGAAATAATGGCAAACAGTTTAGACTACAATGAGAGGCTC
TGGGCTTGGGAAAGCTGGAGATCTGAGGTGCGCAAGCAGCTGAGGCCATTATATGAAGAGTATGTGGTCTTGAAA
AATGAGATGGCAAGAGCAAATCATTTATGAGGACTATGGGGATTATTGGAGAGGAGACTATGAAGTAAATGGGGTA
GATGGCTATGACTACAGCCGCGGCCAGTTGATTGAAGATGTGGAACATACCTTTGAAGAGATTAACCCATTATAT
GAACATCTTCATGCCTATGTGAGGGCAAAGTTGATGAATGCCTATCCTTCCTATATCAGTCCAATTGGATGCCTC
CCTGCTCATTTGCTTGGTGATATGTGGGGTAGATTTTGGACAAATCTGTACTCTTTGACAGTTCCCTTTGGACAG
AAACCAAACATAGATGTTACTGATGCAATGGTGGACCAGGCCCTGGGATGCACAGAGAATATTCAAGGAGGCCGAG
AAGTTCTTTGTATCTGTTGGTCTTCCCTAATATGACTCAAGGATTCTGGGAAAATTCATGCTAACGGACCCAGGA
AATGTTTCAGAAAGCAGTCTGCCATCCCACAGCTTGGGACCTGGGGAAGGGCGACTTCAGGATCCTTATGTGCACA
AAGGTGACAATGGACGACTTCCTGACAGCTCATCATGAGATGGGGCATATCCAGTATGATATGGCATATGCTGCA
CAACCTTTTCTGCTAAGAAATGGAGCTAATGAAGGATTCATGAAGCTGTTGGGGAATCATGTCACTTTCTGCA
GCCACACCTAAGCATTTAAATCCATTGGTCTTCTGTCAACCCGATTTTCAAGAAGACAATGAACAGAAATAAAC
TTCCTGCTCAAACAAGCACTCACGATTTGTTGGGACTCTGCCATTTACTTACATGTTAGAGAAGTGGAGGTGGATG
GTCTTTAAAGGGGAAATTTCCAAAGACCAAGTGGATGAAAAAGTGGTGGGAGATGAAGCGAGAGATAGTTGGGGTG
GTGGAACCTGTGCCCCATGATGAACATACTGTGACCCCGCATCTCTGTTCATGTTTCTGATGATTACTCATTC
ATTGATATTACACAAGGACCCCTTACCAATTCCAGTTTCAAGAAGCACTTTGTCAAGCAGCTAAACATGAAGGC
CCTCTGCACAAATGTGACATCTCAAACCTCTACAGAAGCTGGACAGAACTGTTGTAAGAAATACCTCAAATGTT
GAACCTCTCCTAGTATTCAATATTACTCATTTCCATGCCCTAGGTTTGTATTTGATTTCTTTGTTCTAAAAAGAAA
ATTTTATGGCCTCAAATGTCTCATTTACAAACCAAACATTTAATTTGTGGTCAGACAGGAACCTAGACCATAC
AACAATTGGGTGGGCCACCTCTTTCTCCCTATCATAACTACAGCCCTCTCTTCTGGTAATTGGAAGGAAAGAG
CGGTTTAGGGTGGAAATATATCTGTTAATATGCATTCTTTCTTATCTGCCAGAAGCAAATTTAGCCAAGTCAAAG
AGAAGAAACCATAGATCATAGATGTAATATATGTACATCTGGAACCCCTCAAAGGCCCTGAACCCCTTTTTT
TGTGTAGCAATATGCTGAGGCTTGGAAATCAGAACCCCTGGACCCCTAGCATTTGGAAATGTTGTAGGAGCAAGAA
CATGAATGTAAGGCCACTGCTCAACTACTTTGAGCCCTTATTTACCTGGCTGAAAGACCAGAACAAGAATTTCTTT
TGTGGGATGGAGTACCGACTGGAGTCCATATGCAGACCCAAAGCATCAAAGTGAGGATAAGCCCTAAATCAGCTC
TTGGAGATAAAGCATATGAATGGAACGACAATGAAATGTACCTGTTCCGATCATCTGTTGCATATGCTATGAGGC
AGTACTTTTTAAAGTAAAAAATCAGATGATTTCTTTTGGGGAGGAGGATGTGCGAGTGGCTAATTTGAAACCAA
GAATCTCCTTTAATTTCTTTGTCACTGCACCTAAAAATGTGTCTGATATCATTCCTAGAAGTGAAGTTGAAAAGG
CCATCAGGATGTCCCGAGCCGTATCAATGATGCTTTCCGTCTGAATGACAACAGCCCTAGAGTTTCTGGGGATAC
AGCCAACACTTGGACCTCCTAACAGCCCCCTGTTCCATATGGCTGATTGTTTTTGGAGTTGTGATGGGAGTGA
TAGTGGTTGGCATTTGTCATCCTGATCTTCACTGGGATCAGAGATCGGAAGAAGAAAAATAAAGCAAGAAGTGGAG
AAAATCCTTATGCCTCCATCGATATTAGCAAAGGAGAAAAATAATCCAGGATTCAAAAACACTGATGATGTTGAGA
CCTCCTTTTAGAAAAATCTATGTTTTCTCTTGGGTGATTTTGTGTATGTAATGTTAATTTTATGTTATAG
AAAATATAAGATGATAAAGATATCATTAATGTCAAACCTATGACTCTGTTTCAGAAAAAAATTTGTCCAAAGACA
ACATGGCCAAGGAGAGAGCATCTTCATTGACATTTGCTTTCAGTATTTATTTCTGTCTCTGGATTTGACTTCTGTT
CTGTTTCTTAATAAGGATTTTGTATTAGAGTATATTAGGGAAAGTGTGTATTTGGTCTCACAGGCTGTTTCAGGGA
TAATCTAAATGTAAATGTCTGTTGAATTTCTGAAGTTGAAAACAAGGATATATCATTTGGAGCAAGTGTGGATCT
TGTATGGAATATGGATGGATCACTTGTAAAGACAGTGCCTGGGAACCTGGTGTAGCTGCAAGGATTGAGAATGGCA
TGCATTAGCTCACTTTTCATTTAATCCATTGTCAAGGATGACATGCTTTCTTCACAGTAACTCAGTTCAAGTACTA
TGGTGATTTGCCTACAGTGATGTTTGAATCGATCACTGCTTTCTTCAAGGTGACAGGTCTAAAGAGAGAAGAATC
CAGGGAACAGGTAGAGGACATTTGCTTTTTCACTTCCAAGGTGCTTGATCAACATCTCCCTGACAACACAAAACCTA
GAGCCAGGGGCCCTCCGTGAACCTCCCAGAGCATGCCTGATAGAACTCATTTCTACTGTTCTCTAACTGTGGAGT
GAATGGAATTTCAAACGTATGTTCAACCTCTGAAGTGGGTACCCAGTCTCTTAATCTTTGTATTTGCTCACA
GTGTTTGAGCAGTGTGAGCACAAAGCAGACACTCAATAAATGCTAGATTTACAAAA

72/550

FIGURE 72

MSSSSWLLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQNMN
NAGDKWSAFLKEQSTLAQMYFLQEIQNLTVKQLQLQALQQNGSSVLSEDKSKRLNTILNTMSTI
YSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVVGKQLRPLYEEYVVLKNE
MARANHYEDYGDYWRGDYEVNGVDGYDYSRQGLIEDVEHTFEEIKPLYEHLHAYVRAKLMNAY
PSYISPIGCLPAHLLGDMWGRFWTNLYSLTVFPGQKPNIDVTDAMVDQAWDAQRIKFKEAEKFF
VSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILMCTKVTMDDEFLTAAHHEMGGH
IQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSIGLLSPDFQEDNETEINFLK
QALTIVGTLPFETYMLEKWRWMVFKGEIPKDQWMKKWEMKREIVGVVEPVPHDETYCDPASLF
HVSDDYSFIRYYTRTLYQFQFQEALCQAAKHGEGPLHKCDISNSTEAGQKLL

Important features of the protein:**Signal peptide:**

amino acids 1-17

N-glycosylation sites.

amino acids 53-57, 90-94, 103-107, 322-326, 432-438, 546-550

N-myristoylation sites.

amino acids 260-266, 286-292, 395-401

Cell attachment sequence.

amino acids 204-207

Neutral zinc metallopeptidases, zinc-binding region signature.

amino acids 371-381

73/550

FIGURE 73

CCCACGCGTCCGAGCGGGGTGGACAAGTGGCGTGTGTGCTGCGACCCCGAGGGAAGATGAACG
GGACGCGGAAGTGGTGTACCCTGGTGGACGTGCACCCAGAGGACCAGGCGGCGGGCGGGCAGGA
AGACCTATGCCATGGTGTCCAGCCACTCAGCTGGTCATTCTCTGGCTTCAGAACTGGTGGAGT
CCCATGATGGACATGAGGAGATCATTAAAGGTGTACTTGAAGGGGAGGTCTGGAGACAAGATGA
TTCACGAGAAGAATATTAACCAGCTGAAGAGTGAGGTCCAGTACATCCAGGAGGCCAGGAACT
GCCTACAGAAGCTCCGGGAGGATATAAGTAGCAAGCTTGACAGGAACCTAGGAGATTCTCTCC
ATCGACAGGAGATACAGGTGGTGTCTAGAAAAGCCAAATGGCTTTAGTCAGAGTCCCACAGCCC
TGTACAGCAGCCCACCTGAGGTGGACACCTGTATAAATGAGGATGTTGAGAGCTTGAGGAAGA
CGGTGCAGGACTTGCTGGCCAAGCTTCAGGAGGCCAAGCGGCAACACCAGTCAGACTGTGTGG
CTTTTGAGGTCACTCAGCCGGTACCAGAGGGAAGCAGAACAAAGTAATGTGGCCCTTCAGA
GAGAGGAGGACAGATGTCCAGAGTGATTGGAGAATGTCTGGGGGAATGAAGTTCCTTCCACA
AACACAGCTCAGTTCTTAGCAACAACTGTTTGTCTTTCTACTTGCTCCATCTGCAGCCTACG
CTGCCCTGGCCTCCTGCAGACAGATAGTGGGGTTACCTGGCAAGGCCTGGTGAGAGCCAGTGA
ACCTAAGCTTTGACTGGGTGGCCTTGCTCTTTCTGGGGAGGAGGGAATGTACATTCAGGGAGTA
GCCTTTTGCGGAAAAATTCTCTAGGGCTACAGACAGTCATGTGTGACTTCTCTCTGCTGTGAA
AACTCCCAGAGTCTCTTTAGGGATTTTCCCTAAGGTGTACCACCAGGCACACCTCAGTCTTCT
TGACCCAGAGCCTGAAAACCTGTTTTCACTGGGTTCCACCAGTCCCAGCAAAATCCTCTTTGTA
TTTATTTTGCTAAGTTATTGGTGGTTTTTGCTTACATCTCATGATTGATATAATACCAAAGTTC
TATAGCCTTCTCTTGCAGTATTTGGATTTGCTTGAAACCGGGAAAACTGTTCCCATTAGGCTT
GTTAATGTCAGAGTGACACTATTATGAATCTTTCTCTCCCTTTCCTCTGCCTGTTTCTPCTCT
CTTTCTCCTTCAAACCTGCTCTGCAGCTAAGGAAGGTGAGTCTACTTTCCCTGAGGCTTTGGG
GTCAGAGTATATGTTGTTTGGAGAAAGAGGGCAATCAGGACTCTTCTGGGACCCAGATGAGTT
CTTCACTAGCCCTTCTGAACCCCTTGCTCCATAATTGGTCTTTTATCCTGGCTCTGAATGACC
CTGCAGGTCATCATGGTTTTCTTTTTTATTTGTTTTTTTTTTTTCTGAGACAGAGTCTCACT
CTGTCACCCAGGCTGGAGTGCAGTGGCGGATCTCAGCTCACTGCAACCTCTGCCTCCCGGAT
TTAAGCGATTCTTCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGTGTGCCACCACGCCTG
GCTGATTTTTGTATTTTAGTAGAGATGGGGTTTCACCATACTGGCTAGGCTGGTCTCGAATT
CCTGACCTCAGGTGATCCACCCACCTCGGCTTCCCAAAGTGCTAGGATTATAGGCTTGAGCTA
CTGCGCCCGGCCCATGGTGTTTTTCTTTAGGGCTCTTCCCTACAGCCTTGAGAAGTAGATAGGC
ATCAGAGTATGGTACTATAGGAATCAGAAAAATTCAAACAAATGTGGATTAAGTGTTTAGGC
TCTATGTGGCTCACGCAGCCAGAATCCTTAAGTCTGTGTGTTTCTGTGTCTCAAGACTGGGCT
CACATTCTGGCTTTGTCCATAACAATGCTCTGGGATTTACAGGAGTTCCTCATTGTAAAT
GAGGGGGTCAGAGCAGGTGATATCCATGTTTCTTCCCTTCTGATATTGTGTCTGTGGCATA
TTCTTTGTATGGCGAATTTAATAAATTATATTAATGTGTCA

74/550

FIGURE 74

MNGTRNWCTLVDVHPEDQAAAGRKTYAMVSSHSAGHSLASELVESHGHEEIIKVYLKGRSGD
KMIHEKNINQLKSEVQYIQEARNCLQKLREDISSKLDRLGDSLHRQEIQVVLEKPNGFSQSP
TALYSSPPEVDTCINEDVESLRKTVQDLLAKLQEAKRQHQSDCVAFEVTLSTRYQREAEQSNVA
LQREEDRCPE

Important features of the protein:**Signal peptide:**

amino acids 1-39

N-glycosylation site.

amino acids 2-6

Amidation site.

amino acids 21-25

75/550

FIGURE 75

GCTTGACACATGGCTCCGGAGGCTCCGGTTGCCCATCCGAGCCCCTGCCAGGCTCTAACGTTCCCAACTGACAA
CACCAGTAACTAAATATAGGAGCAGATGGTGGGGACGGGCTGTGCGAGCGGCTCCTTTGCAGAGGTCTCCGGACT
GCAGATAAGGCTCAGGCCCTTTTGTGAGAAGCAGACCAGCCTGGGGGCTGGCGGCAGGACACCTGTGTCTGCATG
CTGAAGAAGATGGGTGAGGCCGTGGCCAGAGTAGCAAGGAAGGTCAACGAGACGGTGGAGAGCGGCTCTGACACT
CTGGACCTGGCCGAGTGCAAGCTGGTCTCCTTTCCCATTTGGCATCTACAAGGTCTGCGGAATGTCTCTGGCCAG
ATCCACCTCATCACCTGGCTAACAACGAGCTTAAGTCCCTCACCAGCAAGTTCATGACCACATTAGTCAGCTC
CGAGAGCTCCACCTGGAGGGGAACCTTCCCTACACCGCCTCCCCAGCGAGGTGAGTGGCCTGCAGCACCTCAAGGCC
ATTGACCTGTCCCGGAACCAAGTTCAGGACTTCCCTGAGCAGCTTACCGCCCTGCCGGGCTGGAGACCATCAAC
CTGGAGGAGAACGAGATCGTAGATGTGCCCGTGGAGAAGCTGGCCGCCATGCCAGCCTTGGCGAGCATCAACCTC
CGCTTCAACCCACTCAACGCCGAGGTGGCGGTGATCGCCCGCCGCTCATCAAGTTTGACATGCTCATGTCTCCG
GAAGGCGCAAGAGCCCCCTACCTTAGGCCACCCCTCCTCATGCCACCCAGCAAGGGACAGAGGCCACAGGCCCTG
GAACCTGGAAGGGAGGGAGGCCCATGGGAGGCCAAGCCTGGGGGCTGGGGGCGGGTGGGCCGAGCAGCAGTGG
TGGGTGGGCTGCAGCTGGTCTGGATAGATAGCTTACAGCAGTAGTGGGCTCTGGAATGCCCAAGGGAAGAGGCCAA
GGTGGGGCCTGCAGCCTGGACTCGGCACCTCACAGCTGCTGTGCAAACTCAGGCAGATCTCCTGCCCTCTCTGAGC
CTTGCTCACTTGAAAAAACAGGACCCCTTCCCTCCTTTGGGCTCCCTGGAGGTTTTTAAGCAGTACGTGCCCTCCA
AGTTACCTCCAGATCAGCAGGCACAGGTGGGCATTGCCAGGTATTTTCTGAGCCCCTGCGGGTTTTGAGGCCCTGT
TTTTAGTGCTGAGAGCCAGTTGCTGCCCTGAGAAGAGAAGACAACCTCCATCTATTTATTGCTTCCTGAGAACTG
ACCTGGATGCGGCCCTCTGCAGGGCCAGTCTTCAGTCTGTGGTCCCTGGACTGGTGGGAACCTGAACTAGGAG
TCCTGGGAGAGCTGTGGTGGGAATATGGGCTGGCACTGCTGCAGGGCAAGAACATTATGTAGGAGCCCGAGGAC
CANCANGCTGGGAATGGGGAGCAAGTCACGTGAGCTCTGTGATTCCCCACAGTTAACAATTTGGCGGGGTGGGAA
GTCTGAGTGCTCCGTCCCTCTAGCATCACTCCTGAGCTGCGGGAGAGGTGGCCAGAGAACAGCAGAGTCACTT
ACACCTGCAGCTCTTGTCTAAAGTGATTAGATGGCCACCCTCACCAGTGTCCAGTCCAGCAGCAGCCTGGCTGCC
TTGTGATGGCCTCCTGGGGGCAGAAAGGCGATGTGGACCACGGGATTTGTAGCCAGCCAGCTCCCAGGCCAACGCC
CAAAGCCCTGATGACCTGGTTCCTCTGAGGCCCTCAACCTGGCATCTTAGGGTATGGTCAGGCAACAGGGTGACC
AGCTGTCTGGTTTTCCAGGACATGGAACCTTCAATGCTAAAACCTGGGACATTAACCCAGCAAGTGGGGATGGTTG
GTCCCCCTACCAGGAGAGGGCCTGGGGCTCTTGCTTCCCGAGAACGCCTGTGGCTTGAAGAACCTTGACTGCTTGG
TCCTCAGGTATCTACCTCCACCTTCTCCTCATCTGTGGAGCAAGCCAACTCAGTCCCCAGACCCACCTGATC
TGCATCTTTGTTTGTCTCCAGAGACACCTGAGGCCCCAGAGCTTGAGGCAAGGCCAGGCCCTCCAAATCCTGTGTG
CCGTGGACGAGTGGCCACTTFACTACTCCTAAGGCTAAGATGTTGAGAGCTCAGACCCTGCTCAGAGCAGTAAT
CCCTGCTCAGAATGCTCCAGTTCCCTCGTCCCTGCCAGGTCTCTTGTCTCTTGGGAAGGAACTGATAGGTCCG
GCCATTGTTGGGCCATCACTGAGCGCTCAGTATCTCAAGAGACTCTGTTCACTCTGCTCGTATCCCAAGGCCTGG
TTGTTCAAACTCTGGGCAAGGCTTTTCAGGATGAGGAGGTCAAGACAGGATGTCCAGAGCTACCGAGTTTCATCT
GTGGGTGTTGGGGGCAAGTGGGGCTGAAGTCCCTGTGCAGGCTGCGCTGGCCCCACCTGCCTTGTGCCCTGGAGT
GGGGTTTCTCCTTGTGAAGAAGAGGCATCCTTCTCTGATGTGCACAAACACAATGTATGACCAGAGCCTTGCAA
CTCAAAGTGTGGTCTGTGGACCAGCAGCGGCAGTGACACCTGGGAGCTTGTAGGAATGCAGAGTCTAGGCCTCA
CCCTATACCTCCCGACTCAGACCTTGCATTTTAGCAAGACCCCCAGCTGATTCTATAAGCACTTTAGAGTTTGA
GAAGCAAGGAACCTAGGCTGGGGATGTCTCCGAGCAGAGGGTGAAGTTTCTCTCAGTTCTCTCCCTGCCACTTCC
AGGGATCTGAGCCTGTGTTTACGCCCTCCTCCCTAACCCACCCTGGGAGACACTTGGCCTGTTAGATTGTTCCAGAG
TCTGCATGGCACTCCTGAAGAAGGGAGTGTGACCTGCAGTCACCAGGAGATGAGGTTAGGTGTGCCAGCCCTC
CAGACCCGGCCTTTCTGGTTAACCCCTGCATGCCAAGCTGCCTGCTGCCCCAGGTCTCACCTCAGGCCTTTGAA
GGGGCAGCTTCTGGAAGTTGTTTCTCTCTGCTTGGAGAGTTTGCCCTTGTCTGTCTTGGAAAGTGTGGGCAGC
CACAGATGCCCCCAATCAGAGCTCACAGTGAAGTGAAGCCCTAAGCTTCAGTCTGCAATAAAGAATGCATTGGTT
TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

76/550

FIGURE 76

MLKKMGEAVARVARKVNETVESGSDTLDLAECKLVSFPIGIYKVLNRNVSGQIHLITLANNELK
SLTSKFMTTFSQLRELHLEGNFLHRLPSEVSALQHLKAIDLSRNQFQDFPEQLTALPALETIN
LEENEIVDVPVEKLAAMPALRSINLRFNPLNAEVRVIAPPLIKFDMLMSPEGARAPLP

Important features of the protein:

N-glycosylation sites.

amino acids 17-21, 47-51

77/550

FIGURE 77

CACCAACAAGCAATCGTTTCATGAGAAAGCCGTGCACCCGCTGCAGTTGGGCCATGTGGTCCGCATCGTATTCCAC
 TAGGTCCCCATTGTACACCAAGTACTGTCCCGCGCTCTCCAGCAGATGCCTGCAGCCTTCCACCTTCTCAAGCAG
 GGTGGTGTGAGTGCCTGCTTCTCTCTCGCCTGGACCGGAGCCGTGCGGGGAGGCACCCCCGGGGTGGAGAA
 AAAGCCGGCCTGGCCTCGGAGGTGCTCTCGCCCCCCCCGCCACCGACTCCCTCCTCCCCCTCCAGAGGCGGGCGGC
 GGCTCCGGCGGCAGCAGCGGCAGGCAGCAACGTAAGCGGGATGCTCTCCAGGCTGCTTTTCTGCTCGGTTCAGCAA
 ATGGCTGAGCTGGTACATCTCGCTCTCCAGGTAGGAGATCTCGCGGGCCGTCTCTATGAAGTGGCCGTAGTTCTG
 GTAGACGTTGCGCTTCAGGTTCTGCGCCGTCTCTCCGCCAGCGCCTGGATGCGCTGCGCGTGTCTCTGGAGGTC
 CCGGTCCCCATCCGACTGCTGCGAGAGCTGCTTACGTACAGCCGCGCCTCAAACCCCCCTGACTCCAGCTGCCG
 ACGCAGGCGGCTCGCCCCACTGTCCGACATCGCCATTCCTCTCCGGGTCTCAGGCACTCACTGTCACTA
 TCGGCGCCGACAGCCGCGCGCTGTCTAGACCCACCCAGGCCAACCGAGCTCCTGGGCTGAGGAAGCAGGAATG
 GGAACGAGACGAGTACGCCTGCGCGGGTCTGAGCGTCAGACACTGCGCCTGCGCAAGTGGGCCGAGCGCAGACA
 TTGCGCCTGCGCAGCAATGCCATCGGTTAAAGCGCATGCGCAAGATGAGCTATTGCGGAAGTGAGGGGAGGGAGA
 GGCCGAGAGAAATTTGCGTACTGCGCATGAACCGAGCGTGACGTTGAGGTTTGAATAACCGGCAAGAGTAAAG
 GCTGAAACTAGCTTCTGAAAGCTTCGTAGGGCCCGAGCCCTGTGAGCCCGAGGTTCTGCGCCCACTAGGAGGTGT
 CATGCTGACTGCTTTTTTAAAGCCCTAGAATCCTTGGCTTCGGCGTTTGGGGTAAGCTCCGTTCTCGTTCTCAA
 GCGCGTTTCCGCGAACTCTCGCGGGATTGACGGGCGCTCTCGAGAGCCGGCATCTCTTAGGAGCTAGTCCTGGTC
 CTCGGCTAGGCGGCTTGGGGTGGCGCGTAACGGGGAGCCAGCCTGACGCCGGCGGACCCCGCTGTGATCCTG
 GCAACGATGATGACTTGATGTTGGCACTGCGGCTTCAGGAGGAGTGGAACCTGCAGGAGCGGAGCGCGAT
 CATGCCCAGGAGTCCCTGTGCTAGTGGACGCGCTGTTGGGAGTTGGTGGACCCACACCGGACTTGCAGGCACTG
 TTTGTTCAGTTTAAACGACCAATCTTCTGGGGCCAGCTGGAGGCGCTCGAGGTGAAGTGGAGCGTGCGAATGACC
 CTGTGTGCTGGGATATGCAGCTATGAAGGGAAGGGTGAATGTGTTCCATCCGTCTCAGCGAACCCCTTTTGAAG
 TTGAGGCCAAGAAAGGATCTTGTAGAGACCCTCTGCATGAAATGATACATGCCTATTTATTTGTCACTAATAAC
 GACAAAGACCCGAGAAGGGCATGGTCCAGAATTTTGTAAACATATGCATCGCATCAACAGCCTGACTGGAGCCAAT
 ATAACGGTATACCATACTTTTACGATGAGGTGGATGAGTATCGGCGACACTGGTGGCGCTGCAATGGGCCGTGC
 CAGCACAGGCCACCGTATTACGGCTATGTCAAACGAGCTACTAACAGGGAACCCCTCTGCTCATGACTATTGGTGG
 GCTGAGCACCAGAAAACCTGTGGAGGCACTTACATAAAAAATCAAGGAACCAGAGAATTACTCAAAAAAGGCAAA
 GGAAGGCCAAAAGTAGGAAAGGAACCACTATTGGCCGAGAGAATAAAGTACCTTCGTGTATATTCTTCTGATT
 TTTATGACCATATGATGTAAAGACAATACGTCCTTCAGAGAACCTGGTATTAAAGATAAACTTAAGGATC
 GTTCTGGTGTAGAAGTCTTCAAGTGTAGACTTAAGGAAAAATCCCACGTGTCATGAATGATGGTAGGAAAC
 AGACTTTGCTCTGTACAGAAGTAAGTAAAGTAGGAATAGTTCCATGGATATTTTTATTTTATTAACCTTTTTT
 CAGTTTCTTTTATTCAAAGAAACAAAATTCATCTCTGATAATATTTGAGGTAAAGTTCCTTTCCCTATCTTGA
 CTCACTGAGTTATTAGGAAACAGAAGGCCAAAAGATTGTCAAATAAAACAATAATTCAAGTAACAATGCCCGG
 AATATACGTCCTAACTACACCCCTTCTATCAGCTGGATTCTATCCAAGTGACTCTATTGATGTATGTATGTTCA
 TTCAAAGAATGGGAAAAGGATATGACATATATTTGCCAGTACTTCATCTTCAAGATTTACCCCTTTTCTGTGAAG
 TTCAGAGTTACTGAAGATGCTTCTTCCCTTGGGAAGTTGTTGACCCAGAACATAGGTTATATTTCCAAATCTT
 TAATTATTGAGTGAAGAGCTATAGATGAATTGATATGGAAGACCGTATCTTCATTTTCGTGAGTAGAAGGAAA
 GATAAGAATGAGGCAGCAGATTTTCCCTCCTGGAATTACACATAAAGGACACTAAGCAATTTTCAAGGTAAATGT
 TGCCTTGTGTGGTCTTTGGCATGATAAGATTCTTTATTTAAATATGAGAGAATTTTTTTTTATCCTTTATATT
 CTCTCAATATCAGAACTCCTGAATTCTGAAGATTGCCCTCCTCCCATTAATAGGATTGTATGGATGTAAGATGGA
 ATAAATACTAGTTCTTCATTTTGAAGAACTGTACATTAGTTTAAATGTTTGTACTGTATTTCTTTGAGTTGA
 GGCATTTACATAACAATCTTCTTTGCTTTTTTGGCAGATAAACCCAACAGAGGTGAGGCCAGCTAGTAATCCCT
 TTTAGTGGGAAAGGATATGTTCTAGGAGAAACAAGCAATTTACCTTCACCTGGGAAACTGATCACTTCACATGCC
 ATTAATAAAACCCAGATCTTTTAAATCAAACCATTCAGCAATGCTGTAAGACCTAATTTCTAAATCAAGGTG
 AATTTGAACAGAATGGTTCAAGTAAAAATTCATCTGGTCTCCCTGCTGTAGTAACAGTCACCAAATGTT
 CTAAGCAACTACTTTCTTAGAGTATCATTTGCCAACCAAAAGGCTTTCAGAGGTGTAATGGATCTCCAAGGATA
 AGTGTAACAGTTGGCAACATCCCTAAAACTCAGTCTCTCTAGTTCTCAGAGAAGGGTTTCATCTTCTAAGATA
 TCCCTAAGAAATTTCTTCAAAGTAACGGAATCAGCATCTGTGATGCCATCCAGGATGTGAGTGGGTCTGAAGAT
 ACATTTCCCAAATAAACGACCTAGGCTAGAAGATAAAAAAAA

78/550

FIGURE 78

MDDDLMLALRLQEEWNLQEAERDHAQESLSLVDASWELVDPTPDLQALFVQFNDQFFWGQLEA
VEVKWSVRMTLCAGICSYEGKGGMCSIRLSEPLLKLRPRKDLVETLLHEMIHAYLFVTNNDKD
REGHGPEFCKHMRINSLTGANITVYHTFHDEVDEYRRHWWRCNGPCQHRPPYYGYVKRATNR
EPSAHDYWWAEHQKTCGGTYIKIKEPENYSKKGKGKAKLGKEPVLAAENKGTFFVYILLIFM

Important features of the protein:**Signal peptide:**

amino acids 1-41

N-glycosylation sites.

amino acids 148-151, 217-220

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 184-187

Casein kinase II phosphorylation sites.

amino acids 30-33, 121-124, 154-157, 187-190, 192-195

Tyrosine kinase phosphorylation site.

amino acids 211-218

N-myristoylation sites.

amino acids 59-64, 85-90, 146-151

Neutral zinc metallopeptidases, zinc-binding region signature.

amino acids 108-117

79/550

FIGURE 79

CGGACGCGTGGGTGGCAACCAGGAGAAGCCAACTTGGTCCCCCGGCTCGCGGAGTGCCTGCG
AGCGGTGCTC**ATG**GCGCTCTATGAGGTCTTCTCTCACCCGGTCGAGCGCAGTTACCGCGCGGG
GCTCTGCTCCAAAGCCGCGCTGTTCCCTGCTGCTGGCCGCTGCGCTCACGTACATCCCGCCGCT
GCTGGTGGCCTTCCGGAGCCACGGGTTTTGGCTGAAGCGGAGCAGCTACGAGGAGCAGCCGAC
CGTGCGCTTCCAACACCAGGTGCTGCTCGTGGCCCTGCTCGGACCCGAAAGCGACGGGTTCCT
CGCCTGGAGCACGTTCCCCGCTTCAACCGGCTGCAAGGGGATCGCCTGCGCGTCCCGCTCGT
TTCGACTAGAGAAGAAGACAGGAACCAGGATGGGAAGACGGACATGTTACATTTTAAGCTGGA
GCTTCCCCTGCAGTCCACGGAGCACGTTCTCGGTGTGCAGCTCATCCTGACTTCTCCTATCG
ATTACACAGGATGGCGACCCCTCGTGATGCAGAGCATGGCGTTTCTCCAGTCCTCCTTTCCTGT
CCCCGGATCCCAGTTATACGTGAACGGAGACCTGAGGCTGCAGCAGAAGCAGCCGCTGAGCTG
TGGTGGCCTAGATGCCCCGATACAACATATCCGTGATCAACGGGACCAGCCCCCTTGCCTATGA
CTACGACCTCACCCATATTGTTGCTGCCTACCAGGAGAGGAACGTTACCACCGTCTCTGAATGA
TCCCAACCCCATCTGGCTGGTGGGCAGGGCCGCAGATGCTCCATTTGTGATTAATGCTATCAT
CCGATACCCTGTGGAAGTCATTTCTTATCAGCCAGGATTCTGGGAGATGGTAAAGTTCGCCCTG
GGTACAGTATGTCAGCATCCTGCTTATCTTCTCTGGGTGTTTGAAAGAATCAAGATCTTCGT
GTTTCAGAATCAGGTGGTGACCACCATTCCTGTGACAGTGACGCCCCGGGGAGACTTGTGTAA
GGAGCACTTATCCT**TAGA**AAAGGCCATTTCTGAAGACTCAGCAGGACCGTGGCTGCCTCATTGTC
ATCTTCTGGGAACATCTTAGGACCTTTTGAAAGAGCCCAGCGGACACCTGCGGGCTTGTGTGC
TTTTCCCTCAGAGACAACGGTTCTTTCCGGTTTTGCTCTACACAGTTCCTGATCTTCAGAGCT
CCTGCAGAATTGTCAGGGACTAGTTTGTGGAAAGGTCTGAGAGTTCCTGGAGGCTATAATTAG
CTTTTTGGGTTTTCTTCTTTGCCTTAGCGTTGAATTTCAGGAGAAAATGTCAGTCAGTTCAG
ACATCTTGGAAGAGTCCCATCTCTGGTCAAGCAGAGACTTTTCCTCTGTTGAACTGAGGAAC
ACACTGTGCATTTCTTCTTCTGTTGTGAGCCACTCTTACTCTTTTCAGGGCTCTCTTGTGAC
AAACATGCCAATCACTAGCACTTTGCACCCCTGGGCTTCTCCATTTCCCATTCACAGCTTTGA
TTTTCCAGAGCTGAGGCCTTTAACTGGAGACCTGGAGGGGCAGGGCCCAAGGGCAAGGGCCGCA
TTAGCACAGGCAATCAGGGAGGGCCGCTGAAGGACACTTGGACCGTCCACCTGCCCCAGCCCA
ACAGTCAGTCATCTGTCATCAGCTCAGCTGAGCAGCCCTGGATCTTTGCCGTA CTGTGACTGG
GCTCTTTGCCCTATTTTTCCCTCTGTCTGTGCCCTGGATGGCAGGCTGAAGTCAGAGGGGCT
GTTTCATTCTCAGCCCCCTCAGCAGCACTGGGGGAAGAAAGCATTGTCACAACAGGTTCTTTC
TGGCCCTCACCCAACAGCCTGGGCACTTGGCCCTCCTCCTCTTGACAGCCCTCCCCCTTCCT
GCAAAGGACAGGGGCGACAGGGGTGGTGTGGGATTGGCTCCCGCTGCCTGACAACCACAAG
TTTATTTGGAAGGCTAGCGGGAAGCCCAGCGCTGGCGTTCCCTTGACTAAGGAACAGGGTG
CCCATCAGAGTGGGGCGGGCAGCTTTGGGAAGGACACAAGAAGCAGTAAGAGTGTAAGAGGA
TGCTGGCCTGGGCAGGCCAGTCCAGCCTGGCCACTAGCAGAATACCAAGCAGTCCAGTGGATT
ACCCTCGTGGCTAAGCAAGTGTCTGCAGGAGCAGAGATGGCTGGAAGGGGCCTCTGCACACGG
AAGATGGCTTGTTACAGCCATTACCTCCTGAGGATGTGGGCAGTCTCCTCCAAGAACACATG
GAGCTGCTTCTGATCCCAAGCAGGTCATTGCCACTGGAAGGACATGGCCCCGGTGATCCATG
CTTCATGCCCAACCAGAAACACACCCCTCAGTGTGTGCCTCAGTTTACTTTGGAGATCAGTTG
TCGTTTTTAGTGCTCCTTTAGGCTTACTAAAACAGTTTTGGAACAAAGCTATTTTGAAGTAT
TCAAGCAGAGGAATTCCCTAACACTGACCCCTTGCTTTTTTTTAAATATTCAGGCTGTTTTAT
ATGCCATAAATTTTTTTCTTAAGATCTAAACGAAAAATAGTTTCTTGTTTAAATTCACATAAGG
CAATGAGATATGGAAAGATGACAAGATACGTATAAACATTGGTTTGCATCTTATTAAATTATT
CTAATGCAAATCTTGATATAAAGAACCCATGATGTTTTGTAACCTTCTAAPTAAATGTTCAA
ATGAG

80/550

FIGURE 80

MALYEVFESH PVERSYRAGLC SKAALFLL LAAALTYIP LLLVAFRSHGFWLKRSS YEEQPTVRF
QHQVLLVALLGPESD GFLAWSTFP AFNRLQGDR LRVPLVSTREEDRNQDGKTDMLHFKLELPL
QSTEHLV LGVQLILTF SYRLHRMATLVMQ SMAFLQSSFPVPGS QLYVNGDLRLQ QKQPLSCGGL
DARYNISVINGTSPFAYDYDLTHIVAAYQERNVTTVLNDENPIWLVGRAADAPFVINAIIRYP
VEVISYQPGFWEMVKFAWVQYVSILLIFLWVFERIKIFVFQNQVVT TIPVTVT PRGDLCKEHL S

Important features of the protein:**Signal peptide:**

amino acids 1-34

Transmembrane domain:

amino acids 268-284

N-glycosylation sites.

amino acids 194-198, 199-203, 221-225

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 51-55

Tyrosine kinase phosphorylation site.

amino acids 250-259

N-myristoylation site.

amino acids 187-193

Cell attachment sequence.

amino acids 307-310

81/550

FIGURE 81

GCCCCGAGCTTCCTCGATGCTGCCGCCGCCCTCCGAGCCGGGGAGGAGCTGCCAGGGGCCAGCTGGGCAGGAGCCT
GGGTCCGCTGCTGCTGCTCCTGGCGTTGGGACACACGTGGACCTACAGAGAGGAGCCGGAGGACGGCGACAGAGA
AATCTGCTCAGAGAGCAAAATCGCGACGACTAAATACCCGTGTCTGAAGTCTTCAGGCGAGCTCACCACATGCTA
CAGGAAAAAGTGCTGCAAAGGATATAAATTTGTTCTTTGGACAATGCATCCAGAAAGATTACGACGTTTGTGCCGA
GGCTCCCTGTGAACAGCAGTGACCGGACAACTTTGGCCGAGTGCTGTGTACTTGTATCCGGGATACCGATATGA
CCGGGAGAGACACCGGAAGCGGGAGAGCCATACTGTCTGGATATTGATGAGTGTGCCAGCAGCAATGGGACGCT
GTGTGCCACATCTGCATCAATACTTGGGCGAGCTACCGCTGCGAGTSCCGGAAGGCTACATCCGGGAAGATGA
TGGGAAGACATGTACCAGGGGAGACAAATATCCCAATGACACTGGCCATGAGAAGTCTGAGAACATGGTGAAAGC
CGGAACCTTGCTGTGCCACATGCAAGGAGTTCTACCAGATGAAGCAGACCGTGCTGCAGCTGAAGCAAAAGATTGC
TCTGCTCCCCAACAAATGCAGCTGACCTGGGCAAGTATATCACTGGTGACAAGGTGCTGGCCTCAAACACCTACCT
TCCAGGACCTCCPGGCCCTGCCTGGGGGCCAGGGCCCTCCCGCTCACCAGGACCAAGGGAAGCCAGGCTTCCC
CGGTATGCCAGGCCCTCCTGGGCGAGCCCGGCCACGGGGCTCAATGGGACCCATGGGACCATCTCCTGATCTGTC
CCACATTAAAGCAAGGCCGGAGGGGCCCTGTGGGTCCACCAGGGGCACCAGGAAGAGATGGTTCTAAGGGGGAGAG
AGGAGCGCCTGGGCCCAGAGGGTCTCCAGGACCCCTGGTTCTTTGACTTCTGCTACTTATGCTGGCTGACAT
CCGCAATGACATCACTGAGCTGCAGGAAAAGGTGTTGGGGCACCAGGACTCACTCTTCAGCAGAGGAGTTCCCTTT
ACCTCAGGAATTTCCAGCTACCCAGAAAGCCATGGACCTGGGCTCTGGAGATGACCATCCAAGAAGAACTGAGAC
AAGAGACTTGAGAGCCCCCAGAGACTTCTACCCATAGCACATCCCAACACCGTCACGCCAAAGGAAGAGAAAGAT
CAACTCACCTGCAGTTAAACCATCTAAAGAGAAAGAAAGACCCTGGAGACCTAGAAAACATACATTTTCTCTTC
TCTTCTCCTGAGCTCTCTCCACTCCTCTCTTCCAAATACGATGCTATTTTCAGAGTCCCTCCTAGGCCTGCAG
ACATGAGGGAGTGAATGATTGATTACCTGCTTCTCACTAAGAGTCCATTGGGGTGGTTTGCATTGTAACCTTTTC
TTTTACATCCTATTTTTCCAGGAACCTTTGGATTAAAGTACTCTCACAGTGTCTTAAATCATAAATTCTTGAAGTT
AAATTTGGCAGAGTATCAAAAGGGGGAAAAATGACAAAGTGAGCTCTAAGAAAATGTGAGGCTACTTCTAAGATGT
GTGTTCACAATAGACCATAACTCCTCTAGTATCAAAATTTGGGGCTCTTCAGTTAAAAGGGGTGGGGAGGACAAA
CGTGTGATGTGCTTTGGTGGAGAATTTTTCTCTGTGCTTCTAGTAGACTTTAAATATTGTATCCCTTTGTCAA
ACCTTGTTTCCCAAATTCATTAAGAGAGGAGAGAATTGAATGGCGTTTAGAGAAGATAGAAAAGAATCACAGT
CATATATTTACTGTTATATAGATTGCCACATTTCTAAATTCAAATACGGTGCTTAAGGTTTCATGCCATGCTTAT
CTGTAAGTATCCTATTTAGGGAAGAAGATTAACTCTCTTTTCAAAAAACAAAGTGAAATGCCTGGATTACAT
TAAAACAATGGECTCTCGTTTGGCTATAATATTTAAAGCTGTTTAAATCAACAGTGGAGTCTGCTCTATAAATATA
GATTATTTGTTCAATAAACTGGCTGAGCTTAGAGAGAGGTGCAGAATTCCTGGTCTGAGCAGGTGCCCAGAAGG
TACCATTAGGTGCCATGATCCAGGCTGAACCAATATACAGTGGGGCTGAAGTCTGCAAGGAGGTTGCTGGCTTGG
GCTGACCTCACTAATGCCATCAGCAGCGGTAGGTAAATTTTTCTCCTTGGGTATTACAAGTTTTTGTCTGGAGC
CAACCAAGCTTGCCACCAACATATTGAGAGTAATACACTATTGAAAGTTATCTTGGATGGGGAGAAAAAAATA
GTGGTTTTCTTGTGTTGCAAAAACCTCCCTTCCTATTTCTCATTTTTTTCTTAATTTCTTTAATTTAGTCCAAGTTC
CAGTTCTTTTAGGCCTTCTCTTTGATTTATTTTCCCTTGCAATGTGAGAAGCAGTTTCAAAAAAGGTCTATATCTC
CACCTCCTAGTGAGTTAGAGTGTTTCTCAGAGCACCTCTGGGTGGCAAGGGAAGCATGTTCTCTGCCAAGGTTT
GCTGTGGATTGAGAAGCACCAGGAGCAAGAGACCAGAAGGATGATCTGCTCCTTTGTAACTGTTTGGGGCCCT
CTTGTTCCTAATGAGCAGCTTATAGGTTACTCACAGTCCACTTTCTCACTGGACACACAAAGTGGCTCTTTATCT
ACCTTTGGGGGAGATTTTCACTCTCCTGCAATGATCGTTCTCACACTCATATTAGCTCATGTTGGAATTTCCCA
TCCTGCCATGTCTTTCCCATTTCTTTTGGCTTTTGTGCTCCACCTTTTAGCCACATCATTTAACTCCACTA
CTGTGAAAGCTTGCTTAAAGAAAATCCCTCTTGGCCGGGTGTGGTAGCCACGCCTCTAATCCAGCACTTTGGG
AGGCTGAGGCGGGGAGATCACAAGGTGAGGAGATCGAGACCAGCCTGACCAACATGGTGAACCCCTGTCTCTACT
AAAAATACAAAATTAGCTGGGCGTGTGGGCACACACCTGTAAATCCAGCTACTCAGGAGGCTGAGGCAGGAGAA
TTACTTTAACCTGCGGGGGGAGCCTAGATTTGCGCTACTGCCTCCAGCCTAGGCAACAGAGGGAGACTCTGTCTC
ATTAAAAA

82/550

FIGURE 82

MVPPPPSRGGAARGQLGRSLGPELLLLLLALGHTWTYREEPEDGDREICSESKIATTKYPCLKSS
GELTTCYRKKCKCKGYKFVLGQCIPEDYDVCAEAPCEQQCTDNFGRVLCCTCPGYRYDRERHRK
REKPYCLDIDECASSNGTLCALHICINTLGSYRCECREGYIREDDGKTCTRGDKYPNDTGHEKS
ENMVKAGTCCATCKEFYQMKQTVLQLKQKIALLPNNAADLGKYITGDKVLASNTYLPGPPGLP
GGQGPPGSPGPKGSPGFPGMFPGPPGQPGPRGSMGPMGSPDLSHIKQGRRGVPVGPPGAPGRDG
SKGERGAPGPRGSPGPPGSFDFLLMLADIRNDITELQEKVFGHRTHSSAEFFPLPQEFFPSYP
EAMDLGSGDDHPRRTETRDRLRAPRDFYP

Important features of the protein:**Signal peptide:**

amino acids 1-34

N-glycosylation sites.

amino acids 142-148, 182-188

Tyrosine kinase phosphorylation site.

amino acids 125-132

N-myristoylation sites.

amino acids 10-16, 143-149, 155-161, 196-202, 250-256

Amidation site.

amino acids 299-303

Aspartic acid and asparagine hydroxylation site.

amino acids 150-162

Cell attachment sequence.

amino acids 176-179

Clq domain proteins.

amino acids 247-280

Calcium-binding EGF-like domain proteins pattern proteins.

amino acids 144-165

83/550

FIGURE 83

ATCTGAGTGAGCTAACTGACACAATGAAACTGTCAGGCATGTTTCTGCTCCTCTCTCTGGCTC
TTTTCTGCTTTTTTAACAGGTGTCTTCAGTCAGGGAGGACAGGTTGACTGTGGTGAGTTCCAGG
ACCCCAAGGTCTACTGCACTCGGGAATCTAACCACACTGTGGCTCTGATGGCCAGACATATG
GCAATAAATGTGCCTTCTGTAAGGCCATAGTGAAAAGTGGTGGAAAGATTAGCCTAAAGCATC
CTGGAAAATGCTGAGTTAAAGCCAATGTTTCTTGGTGACTTGCCAGCTTTTGCAGCCTTCTTT
TCTCACTTCTGCTTATACTTTTGCTGGTGGATTCCTTTAATTCATAAAGACATACCTACTCTG
CCTGGGTCTTGAGGAGTTCAATGTATGTCTATTTCTCTTGATTCACTTGTCATAAAGTACATTC
TGCAAAAGCAAAAA

84/550

FIGURE 84

MKLSGMFLLLSLALFCFLTGVSQGGQVDCGEFQDPKVYCTRESNP HCGSDGQTYGNKCAFCK
AIVKSGGKISLKHPGKC

Important features of the protein:

Signal peptide:

amino acids 1-23

N-myristoylation sites.

amino acids 26-32, 52-58, 56-62, 69-75

Kazal serine protease inhibitors family signature.

amino acids 40-63

FIGURE 85

GGAGCAGACACACAGACCCCGGGCCGGAGGCCCTCTTCTAGCCCTGCGGGAACCGGACAGTTCC
CCCAAGACTGGGACTCTGGAACACAGCTCCTAAATCATCAAATTTCTCAAGCTTTTTTTTTTCCC
TCTCTTCGTCCCAGCCATCCCAGTCTTCTTCTTCTTTTTTTTTTTTTTAACTTATTGTTTTTT
TCGCTCCTGTCATTATGAAAGTGGTCACGCCATTCAATATTAAGACTTGGAGGGAATTGGGGA
AAGAAAAGAAAGAATCTAAAAGAAGAGAAGCGACCGGTGCTTTTAAGGGTGTCTAATTTTCAA
AAGAGACGTCTGGGAGTATTTTGCTCTGGGCGTTTGGAGCAACTTCGCGGACAGCGGAGCTCG
CCCAGCATGGATGTTCCAGGTTACAGGCGCCTTTCTTCTGAGAACGACCTTGGCCTTGAACG
TCAGAGCCGGGGACGAAGGCCCCCGGAGGCTGCTGCGAGCTCCGCGCGTTCCTTCGCGCCCTT
CCGCGCCGCTCGCGCCGGCGCCGGCCTCCACCCCCGCGCGCCGCTCCCACCAGTCCCGATGC
AGGCGCCCCGGCCGGGGGCCACTCGGGCTGCGGCTGATGATGCCCGGGCGCCGGGGGGCGCTGC
GCGAGCCTGGCGGCTGCGGATCCTGCCTGGGGGTGGCGCTGGCCCTGCTGTTGCTGCTACTGC
CCGCCTGCTGCCCCGTGCGGGCGCAGAACGACACGGAGCCCATCGTGCTGGAGGGCAAGTGCC
TGGTGGTGTGCGACTCCAGCCCGTGGGCGGACGGCGCCGTACCTCCTCCCTAGGCATCTCCG
TGCGCTCCGGCAGCGCCAAGGTGGCCTTCTCCGCCACGCGGAGCACCAACCACGAGCCGTCCG
AGATGAGCAACCGCACCATGACCATCTATTTTCGACCAGGTATTAGTAAATATTGGCAACCACT
TTGATCTTGCTTCCAGTATATTTGTAGCACCAGAGAAAAGGGATTTATAGCTTCAGCTTCACG
TGGTCAAAGTGTATAACAGACAAACCATCCAGGTCAGTTTAATGCAGAATGGCTACCCAGTGA
TCTCGGCCTTTGCAGGAGACCAGGATGTACCAGAGAAGCTGCTAGCAATGGCGTGTCTGCTGC
TCATGGAAAGGGAAGACAAAGTGCATCTCAAACCTTGAGAGAGGCAACCTCATGGGGGGCTGGA
AATACTCCACATTCTCGGGCTTCTTGGTGTTCCTCTATAAAACACAGAGCCCCCTAGATGGTG
GGGGAATGGCAAACCTGGACCCAGGACTCCGCCCTTTAAAACACCCTGAACCTTACTGGAATTGG
ACACCTTGTTTCCAACCTCCGTCAGACTGTTGCAGTAGAAGAATGATTTCTTTGAAACCTCC
AGTACTTTTGTTTTTTGTTTTTTGGGAATACTGACAATTCCTCGGGAACCTGGCCTCTAATTAGT
TTTAGATGACAAGGTCTTAAGGAGAAATGAAATTATCGATTTGAGCAATTTGTACCTGTGATT
GTAAAGTCAATATCGGATTTTATTGTTGGGACCATGGACCTCTTTTGTGTATGTTGTATTG
TCGTCCCAACGGAAGGAGAGCTCCTGACTCCAGGATGGGCTGCAGGTTGCAGTCAGGGCTTGA
AGTAGGAGCCCAGCAAAGAACCACCTGCTGGACAGTCCTTGACATGTGTTCTGTGTGTGTCTG
TATAGCCTTAAGAAAAAGAATGGCTTCACTTTCATTCTGTATTCTTCCCCCACCATGTGGCT
GGGAGGACTTGGGAGGGGGATGGGGACATTGGGAACCTGTCAAGAAGTGCTTTATCCAGAGAA
GCAAAATTTTGACGATTGGACTGCAATTTTTTGTTTTGTATTGTTTGTGTTTTTCTTGAAAAG
CTTTACTTTTCTTTCCACACTCAGCTCTCCCTCCTCAACCCCACTTTTATTTTTCTTGCTGGG
GTTGAGGAGAGAAAAATATAGAAATTCCTGGATAAGACCAAACAAAACAAACATTAAAATACCT
GTATGTTTTGTTTTAGACGAGACCAAACCTAAACAAAAGTATCTGTTTATCAAAGTAAAAGTA
ACACAATGGACAATTCTGCTTATTCTCTCAAAGAGATTCTAAGATGCACCTTTAGAACTATTA
ATAGCAACCTGCATTTTTTTTTTAATTTATACTTCAGAATCCTTTAAGAACCTGGTGTTCCCTGA
GTGETCCTGAATCATATAAGTTGGTAATGGAAGCTGTAATGACCAAGTCCCTTAAACATACTA
TGTCTTTGCCACGTGTGCTGTGACTTCTCTGTGGGTGATTTAATTTATTTGGATCCACCTCTG
AGTGAGCGCACAGTGATCAGGTGCTTCAAAGCCAACAGACCAGCTCCTCTTCCCTCCGGATCCT
CTTTTGATCTGCCCAGGAAAGGGATGCATTGACACTCTCCTGCATGCACCTGGCGAGAAGCCA
CCTGAAAGTCACTGTGGTTAAAGATATTGGTGGAGGTACCCAGGAGCACTGTTACAAATCCT
TCTTGTTTTGGCATCTCGTACAACATTATTAAGACACAGCTGAGAGTTGATGGGTGTGTAATG
CATATGCCAAGGAAATGTCACTAATCCCAAAGCAATCAAAAAGGAGACCTCAAACCAGATGTT
AATTTGTTCTTTGTGTAACAATGTAACCAAAATATTGATGATAAAAGTCATAATTTAAGATTCA
AGAATAAATGGGTTTTGATGTCTGGCAAAAAAAAAAAAAAAAAA

86/550

FIGURE 86

MQAPGRGFLGLRLMMPGRRGALREPGGCGSCLGVALALLLLLLLPACCPVRAQNDTEPIVLEGK
CLVVCDSPPSADGAVTSSLGISVRSGSAKVAFSATRSTNHEPSEMSNRTMTIYFDQVLVNIGN
HFDLASSIFVAPRKGIIYSFSFHVVKVYNRQTIQVSLMQNGYPVISAFAGDQDVTREAAASNGVL
LLMEREDKVHLKLERGNLMGGWKYSTFSGFELVFPL

Important features of the protein:**Signal peptide:**

amino acids 1-48

N-glycosylation sites.

amino acids 53-57, 110-114

N-myristoylation sites.

amino acids 26-32, 27-33, 29-35, 33-39, 76-82, 205-211

Amidation site.

amino acids 16-20

Clq domain signature.

amino acids 117-148

Clq domain proteins.

amino acids 115-149

87/550

FIGURE 87

AGGGCCCCGCGGGTGGAGAGAGCGACGCCCCGAGGGG**ATG**GCGGCAGCGTCCCGGAGCGCCTCTG
GCTGGGCGCTACTGCTGCTGGTGGCACTTTGGCAGCAGCGCGCGGCCGCTCCGGCGTCTTCC
AGCTGCAGCTGCAGGAGTTCATCAACGAGCGCGGCGTACTGGCCAGTGGGCGGCCTTGCGAGC
CCGGCTGCCGGACTTTCTTCCGCGTCTGCCTTAAGCACTTCCAGGCGGTCTGTCTCGCCCGGAC
CCTGCACCTTCGGGACCGTCTCCACGCCGGTATTGGGCACCAACTCCTTCGCTGTCCGGGACG
ACAGTAGCGGCGGGGGGGCGCAACCCTCTCCAACCTGCCCTTCAATTTACCTGGCCGGGTACCT
TCTCGCTCATCATCGAAGCTTGGCACGCGCCAGGAGACGACCTGCGGCCAGAGGCCTTGCCAC
CAGATGCACTCATCAGCAAGATCGCCATCCAGGGCTCCCTAGCTGTGGGTGAGAACTGGTTAT
TGGATGAGCAAACCAGCACCCCTCACAGGCTGCGCTACTCTTACCGGGTCATCTGCAGTGACA
ACTACTATGGAGACAACTGCTCCCGCCTGTGCAAGAAGCGCAATGACCACTTCGGCCACTATG
TGTGCCAGCCAGATGGCAACTTGTCTGCTGCCCGGTTGGACTGGGGAATATTGCCAACAGC
CTATCTGTCTTTCGGGCTGTCTGAACAGAATGGCTACTGCAGCAAGCCAGCAGAGTGCCCTCT
GCCGCCCAGGCTGGCAGGGGCCGGCTGTGTAACGAATGCATCCCCCACAATGGCTGTGCCCACG
GCACCTGCAGCACTCCCTGGCAATGTACTTGTGATGAGGGCTGGGGAGGCCTGTTTTGTGACC
AAGATCTCAACTACTGCACCCACCACTCCCCATGCAAGAATGGGGCAACGTGCTCCAACAGTG
GGCAGCGAAGCTACACCTGCACCTGTGCGCCAGGCTACACTGGTGTGGACTGTGAGCTGGAGC
TCAGCGAGTGTGACAGCAACCCCTGTGCAATGGAGGCAGCTGTAAGGACCAGGAGGATGGCT
ACCACTGCCTGTGTCTCCGGGCTACTATGGCCTGCACTGTGAACACAGCACCTTGAGCTGCG
CCGACTCCCCCTGCTTCAATGGGGGCTCCTGCCGGGAGCGCAACCAGGGGGCCAACTATGCTT
GTGAATGTCCCCCAACTTCACCGGCTCCAACCTGCGAGAAGAAAGTGGACAGGTGCACCAGCA
ACCCCTGTGCCAACGGGGGACAGTGCCTGAACCGAGGTCCAAGCCGCATGTGCCGCTGCCGTC
CTGGATTACGGGACCTACTGTGAACCTCCACGTCAGCGACTGTGCCCGTAACCCCTGCGCCC
ACGGTGGCACTTGCCATGACCTGGAGAATGGGCTCATGTGCACCTGCCCTGCCGGCTTCTCTG
GCCGACGCTGTGAGGTGCGGACATCCATCGATGCCTGTGCCTCGAGTCCCTGCTTCAACAGGG
CCACCTGCTACACCGACCTCTCCACAGACACCTTTGTGTGCAACTGCCCTTATGGCTTTGTGG
GCAGCCGCTGCGAGTTCCCCGTGGGCTTGCCGCCAGCTTCCCCTGGGTGGCCGTCTCGCTGG
GTGTGGGGCTGGCAGTGCTGCTGGTACTGCTGGGCATGGTGGCAGTGGCTGTGCGGCAGCTGC
GGCTTCGACGGCCGGACGACGGCAGCAGGGAAGCCATGAACAACCTTGTGGACTTCCAGAAGG
ACAACCTGATTCTGCGGCCAGCTTAAAAACACAAACCAGAAGAAGGAGCTGGAAGTGGACT
GTGGCCTGGACAAGTCCAACCTGTGGCAAACAGCAAAACCACACATTGGACTATAATCTGGCCC
CAGGGCCCCCTGGGGCGGGGGACCATGCCAGGAAAGTTTCCCCACAGTGACAAGAGCTTAGGAG
AGAAGGCGCCACTGCGGTTACACAGTGAAAAGCCAGAGTGTGCGATATCAGCGATATGCTCCC
CCAGGGACTCCATGTACCAGTCTGTGTGTTTGATATCAGAGGAGAGGAATGAATGTGTCAATTG
CCACGGAGGT**TAAGG**CAGGAGCCTACCTGGACATCCCTGCTCAGCCCCGCGGCTGGACCTTC
CTTCTGCATTGTTTACA